

THESIS / THÈSE

DOCTEUR EN SCIENCES

Approche structurale et fonctionnelle de la première enzyme de la voie de la sérine PHGDH, une cible potentielle pour le traitement du cancer

Themans, Quentin

Award date:
2019

Awarding institution:
Université de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**UNIVERSITÉ
DE NAMUR**

FACULTÉ
DES SCIENCES

Département de Chimie

Approche structurale et fonctionnelle de la première enzyme de la voie de la sérine : PHGDH, une cible potentielle pour le traitement du cancer

Dissertation présentée par

Quentin Thémans

en vue de l'obtention du grade de Docteur en
Sciences

Composition du Jury :

Dr Ute Röhrig (SIB Swiss Institute of Bioinformatics)

Dr. Judith Unterlass (*In vitro* Biology - Evotec)

Dr. Serena Sirigu (Synchrotron SOLEIL)

Dr. Laurence Leherte (Université de Namur)

Pr. Johan Wouters (Université de Namur, promoteur)

2019

Graphisme de couverture : © Presses universitaires de Namur

© Presses universitaires de Namur & Quentin Thémans
Rempart de la Vierge, 13
B - 5000 Namur (Belgique)

Toute reproduction d'un extrait quelconque de ce livre, hors des limites
restrictives prévues par la loi,
par quelque procédé que ce soit, et notamment par photocopie ou scanner,
est strictement interdite pour tous pays.

Imprimé en Belgique
ISBN : 978-2-39029-080-3
Dépôt légal: D/2019/1881/31

Université de Namur
FACULTÉ DES SCIENCES

Rue de Bruxelles 61 – 5000 Namur, Belgique

**Structural and functional approach of the first enzyme of the serine synthesis pathway:
PHGDH a potential target for cancer treatment**

By Quentin Thémans

Abstract

Since 2011, deregulating cellular energetics has been identified as an hallmark of cancer. The Warburg effect and its implication in cancer cells are increasingly studied. The serine biosynthesis pathway and its first enzyme, phosphoglycerate dehydrogenase (PHGDH), has been identified as potential therapeutic target. Since 2016, an increasing number of inhibitors targeting PHGDH are published. Up to now, for most inhibitors, we are lacking information of their binding mode.

In the first part of this work, we tried to crystallize the wild type form of PHGDH with fusion protein (MBP) and mutation on the surface of the protein (surface entropy reduction). Only the truncated form of PHGDH leads to a crystallographic structure showing that ASB and ACT domains are essential for the proper assembling of the protein under its tetrameric form. In the second part of this work, we generated a tetrameric model of the full length PHGDH. The assembly of the tetramer and the global folding of the monomer were confirmed by cross-linking coupled with mass spectroscopy experimentations. On one hand, this model helped us to highlight the binding mode of the substrate in the active site. We also showed the importance of R134 and flexible loop containing it. On the other hand, we screened virtually the binding site of substrate and cofactor in order to find new ligands. The screening showed the preferential binding sites in the pocket. L-cysteine and fluoropyruvate have been identified as potential ligands and their binding in the active site was highlighted by crystallography.

At the end of the thesis, we were able to generate the first structure of the tetrameric form of PHGDH, identified new ligands of the active site and the essentials interactions of binding pockets and highlighted the importance of R134 and the loop containing it for the proper activity of the enzyme.

Ph.D. thesis in Sciences

XXX

Laboratoire de Chimie Biologique Structurale (Supervisors : Prof. Johan Wouters)

Université de Namur
FACULTÉ DES SCIENCES

Rue de Bruxelles 61 – 5000 Namur, Belgique

**Approche structurale et fonctionnelle de la première enzyme de la voie de la sérine :
PHGDH, une cible potentielle pour le traitement du cancer**

Par Quentin Thémans

Résumé

Depuis 2011, le dérèglement du métabolisme énergétique a été identifié comme un *hallmark* du cancer. L'effet Warburg et son implication dans les cellules cancéreuses est par conséquent de plus en plus étudié. La voie de la sérine et la première enzyme de celle-ci, la phosphoglycérate déshydrogénase (PHGDH) ont été identifiées comme des cibles thérapeutiques potentielles. Depuis 2016, un nombre croissant d'inhibiteurs ciblant PHGDH sont publiés mais jusqu'à présent peu d'informations quant à leur mode de liaison sont disponibles.

Dans la première partie de ce travail, nous avons tenté de cristalliser la forme sauvage de l'enzyme à l'aide d'une protéine de fusion (MBP) et de mutations en surface de la protéine. Seule la forme tronquée de l'enzyme a permis d'obtenir une structure cristallographie montrant que les domaines ASB et ACT semblent essentiels à l'assemblage de l'enzyme sous sa forme tétramérique. Dans la seconde partie de ce travail, nous avons généré un modèle de l'enzyme sous sa forme tétramérique. L'agencement des monomères et la structure globale de PHGDH ont été confirmés avec des expérimentations de *cross-linking* couplées à la spectrométrie de masse. Ce modèle a permis d'une part de mettre en évidence le mode de liaison de substrat. Nous avons pu montrer l'importance de R134 et de la boucle flexible qui la contient. D'autre part, nous avons criblé les sites de liaison du cofacteur et du substrat afin de trouver de nouveaux ligands de ces sites. Ce criblage a permis de mettre en évidence les sites d'interactions préférentiels au sein des poches. La L-cystéine et le fluoropyruvate ont été identifiés comme des ligands potentiels et leur liaison au sein du site actif a été confirmée par cristallographie. Au terme de cette thèse de doctorat nous avons pu : générer la première structure de PHGDH sous sa forme tétramérique ; identifier de nouveaux ligands du site actif et les interactions essentiels des deux poches de liaison et mettre en évidence l'importance de la boucle flexible contenant R134 sur l'activité de l'enzyme.

Dissertation doctorale en Sciences

XXX

Laboratoire de Chimie Biologique Structurale (Promoteurs : Prof. Johan Wouters)

Je tiens avant tout à remercier mon promoteur le prof. Johan Wouters pour m'avoir accueilli au sein de son laboratoire, avoir cru en moi et m'avoir soutenu tout au long de mon projet.

Je remercie aussi les membres de mon jury pour avoir accepté d'évaluer mon travail et pour leurs remarques constructives vis-à-vis de celui-ci : le Dr. Ute Röhrig, le Dr Serena Sirigu, le Dr Judith Unterlass et le Dr. Laurence Leherte.

Je remercie aussi tout les membres du CBS (anciens et nouveaux) pour le temps passé ensemble au laboratoire, les conversations scientifiques, les moments de rire... Je pense particulièrement à Bernadette pour tous les conseils de jardinage, Grégoire qui m'a appris tout ce que je sais, Jenny pour nos pauses fruits et légumes, Andréa mon comparse depuis le premier jour et tous les autres : Thomas, Sébastien, Jean, Annaëlle, Mégane, ... Je n'oublie pas non plus les stagiaires et les mémorants que j'ai pu suivre au cours de ce travail qui ont chacun à leur façon ajouter une pierre à ce travail de thèse.

Je pense aussi à l'Unité de Recherche en Biologie des Microorganismes (URBM), mon unité d'accueil, pour m'avoir donné accès à leurs équipements et donné de nombreux conseils pour toute la partie biologie moléculaire de ce travail. Mais surtout pour tous les moments incroyables que nous avons passés. Pour ne citer que ceux-là, je n'oublierai jamais ces « beer hour » du vendredi soir, véritable bol d'air dans un projet qui n'a pas toujours été évident. Je ne citerai personne au risque d'oublier des gens mais je suis sûr que vous vous souviendrez de ces magnifiques moments.

Je remercie aussi les informaticiens : Fred et Laurent qui ne me verront plus débarquer dans leur bureau pour une énième requête et ensuite discuter de tout et de rien.

J'ai aussi eu la chance avec ce projet de collaborer avec des gens extraordinaires qui n'ont jamais hésité à prendre de leur temps pour partager leurs connaissances. Je pense à Judith qui m'a accueilli durant deux semaines à l'Institut Karolinska et qui m'a permis de débloquer la partie structurale de ce travail ! Je pense aussi à Serena sans qui aucune des structures cristallographiques présentées dans ce travail n'auraient vu le jour. Le Dr. Alexander Leitner qui a toujours été disponible pour discuter des résultats de cross-linking et Marc pour les expérimentations de spectrométrie de masse et les discussions quant à l'analyse des résultats. Et finalement le Dr. Edouard Mullarky qui même à l'autre bout du monde n'a pas hésité à tester nos composés et partager ses résultats. Les échanges que nous avons eus, ont été tellement enrichissant et ça a été un réel plaisir de voir un tel soutien venant du monde de la recherche.

Je remercie le Télévie pour avoir financé ce projet et surtout Bruno et Aurélie grâce à qui toutes les activités que nous avons pu faire dans le cadre du Télévie ont été d'excellents moments.

Je remercie aussi mes amis qui m'ont toujours rappelé qu'il y avait plein de choses en dehors de la thèse ! Arnaud et Kévin et notre fameux groupe à 6 ! Gwenn pour les soirées à courir derrière (ou sur) des dinosaures. Aurore, Pierre, Pauline, Vincent, Aurélie, Jean, Orian, ... pour ces barbecues jusqu'aux petites heures autour du brasero.

Et finalement j'aimerais remercier mes parents, Fabienne et Daniel, qui même s'ils n'ont jamais vraiment compris tout ce que je faisais, n'ont jamais cessé d'essayer de comprendre et m'ont toujours supporté et soutenu dans tous mes projets. Merci pour tout, vous êtes extraordinaires et je n'aurais pas pu accomplir tout ça sans tout ce que vous avez fait pour moi. J'ai eu une chance extraordinaire de vous avoir. Merci aussi à Ralphou, toujours dans son monde mais qui sera toujours là en cas de besoin !

Et Mahaut qui a dû (eu la chance de ?) me supporter au quotidien, dans les moments de doutes mais aussi dans les moments de bonheurs. Et qui a toujours su être là pour moi, me soutenir à tout moment. Tu sais tout ce que je pense de toi. Merci d'être là.

Liste des abréviations

a.a. : acide aminé

3-PG : 3-phosphoglycérate

ADN : Acide DésoxyriboNucléique

ARN : Acide RiboNucléique

ATP : Adénosine TriPhosphate

BS3 : Bissulfosuccinimidyl suberate

COSMIC : *Catalogue Of Somatic Mutations In Cancer*

FDA: *Food and Drug Administration*

GMQE: *Global Model Quality Estimation*

IC₅₀ : Concentration inhibitrice à 50%

IMAC: *Immobilized Metal Affinity Chromatographie*

IPTG: isopropyl β -D-1-thiogalactopyranoside

LB: *Lysogeny Broth*

MBP: *Maltose Binding Protein*

MTHFD1: méthylènetétrahydrofolate déshydrogénase

NAD⁺ (NADH): nicotinamide adénine dinucleotide

NBD: *Nucleotide Binding Domain*

OIS: *Oncogene-Induced Senescence*

PDB: *Protein Data Bank*

PHGDH : phosphoglycérate déshydrogénase

Ppyr : phosphohydroxypyruvate

PSAT1 : phosphosérine aminotransférase

PSer : phosphosérine

PSP : phosphosérine phosphatase

SAM : S-adénosyl methionine

SBD : *Substrate Binding Domain*

SER : *Surface Entropy Reduction*

SHMT1 : sérine hydroxyméthyle transférase

SPR : *Surface Plasmon Resonance*

TCA : cycle de l'acide citrique

XL-MS : Spectrométrie de masse couplée au cross-linking

VIH : virus de l'immunodéficience humaine

Table des matières

1	Introduction.....	11
1.1	La dérégulation du métabolisme énergétique cellulaire.....	13
1.1.1	Effet Warburg	13
1.2	La voie de la sérine.....	16
1.2.1	Généralités	16
1.2.2	Biosynthèse de la sérine par les cellules saines.....	17
1.2.3	Implication de la sérine dans les voies métaboliques.....	18
1.2.4	La voie de la sérine dans les mécanismes cancéreux	19
1.3	La phosphoglycérate déhydrogénase (PHGDH).....	20
1.3.1	Informations générales sur PHGDH.....	21
1.3.2	Structure de l'enzyme.....	23
1.3.3	Inhibiteurs ciblant PHGDH	27
2	Objectif et stratégie	33
3	Expression, purification et cristallisation de PHGDH.....	37
3.1	PHGDH sous sa forme sauvage	37
3.2	PHGDH fusionné à une Maltose Binding Protein (MBP)	43
3.3	Réduction d'entropie de surface sur l'enzyme PHGDH complète	46
3.4	Expression et purification de l'enzyme tronquée (3-338).....	49
3.5	Discussion	55
4	Modèle tétramérique de l'enzyme PHGDH complète.....	61
4.1	Génération du modèle	61
4.2	Validation expérimentale du modèle obtenu.....	68
4.2.1	Principe et conception expérimentale.....	68

4.2.2	Résultats obtenus	70
4.3	Comparaison du modèle aux données existantes	78
4.4	Conclusion et perspectives	80
5	Etude du site actif de PHGDH	85
5.1	Prédiction de la conformation du substrat au sein du site actif	85
5.1.1	Validation de la méthode de docking	85
5.1.2	Liaison du substrat au sein de la poche catalytique.....	87
5.2	Interactions observées au sein du site actif de hPHGDH.....	90
5.2.1	Comparaison des modes de liaison de ligands connus.....	90
5.2.2	Conservation des acides aminés interagissant avec le 3-phosphoglycérate.	91
5.3	Arginine 134 : Etude structurale et fonctionnelle	93
5.3.1	Effet de la mutation R134A sur l'activité de l'enzyme.....	93
5.3.2	Effet de la mutation R134A sur la structure de l'enzyme et la liaison au substrat	95
5.3.3	Rôle de la boucle comportant l'arginine 134	98
5.4	Conclusion.....	99
6	Criblage de composés au sein du site actif	103
6.1	Génération des ensembles de molécules	103
6.2	Poche de liaison au cofacteur	104
6.2.1	Résultats obtenus pour l'ensemble « drug »	104
6.2.2	Résultats pour l'ensemble « druglike »	107
6.2.3	Discussion.....	112
6.3	Poche de liaison au substrat	115
6.3.1	Résultats obtenus pour l'ensemble « drug »	115
6.3.2	Résultats obtenus pour l'ensemble « druglike »	118

6.3.3	Discussion.....	121
6.4	L-cystéine et fluoropyruvate : deux ligands de la poche de liaison au substrat	124
6.4.1	Cristallisation du fluoropyruvate au sein du site actif.....	125
6.4.2	Cristallisation de la L-cystéine au sein du site actif	127
6.5	Conclusion.....	130
7	Conclusion	137
7.1	Discussion	137
7.2	Perspectives	141
8	Matériel et méthodes.....	145
8.1.1	Thermoporation et expression des enzymes.....	145
8.1.2	PHGDH complète et mutation de surface	145
8.1.3	PHGDH fusionnée à une MBP.....	147
8.1.4	sPHGDH (3-338).....	147
8.1.5	Essais de cristallisation.....	149
8.1.6	Collecte et traitement des données cristallographiques.....	150
8.1.7	Tests enzymatiques.....	151
8.1.8	Génération du modèle.....	151
8.1.9	Couplage des états d'oligomérisation.....	151
8.1.10	Digestion de l'enzyme et acquisition des données par spectrométrie de masse.	152
8.1.11	Génération de la position de ligands au sein du site actif	153
9	Références.....	157
10	Annexes.....	169

1. Introduction

1 Introduction

Bien que d'importants efforts soient mis en œuvre afin de découvrir et développer de nouveaux traitements anti-cancéreux (et la mise en place de ceux-ci), le cancer est toujours une des causes principales de mortalité dans le monde. En 2018, 9,6 millions de personnes sont mortes d'un cancer¹. Une mort sur 6 dans le monde se produit à la suite d'un cancer. En Belgique, 68.216 nouveaux cas de cancer ont été enregistrés, et on estime que 1 homme sur 3 et 1 femme sur 4 auront un cancer avant leur 75 ans².

Les cellules cancéreuses ont été observées pour la première fois en 1831 par Virchow and John Hughes Bennett³. Les recherches contre le cancer ont réellement débuté en 1971 avec la mise en place du « National Cancer Act of 1971 » aux Etats-Unis⁴. Depuis lors, les recherches s'intensifient et de plus en plus de traitements commencent à être disponibles (Figure 1-1).

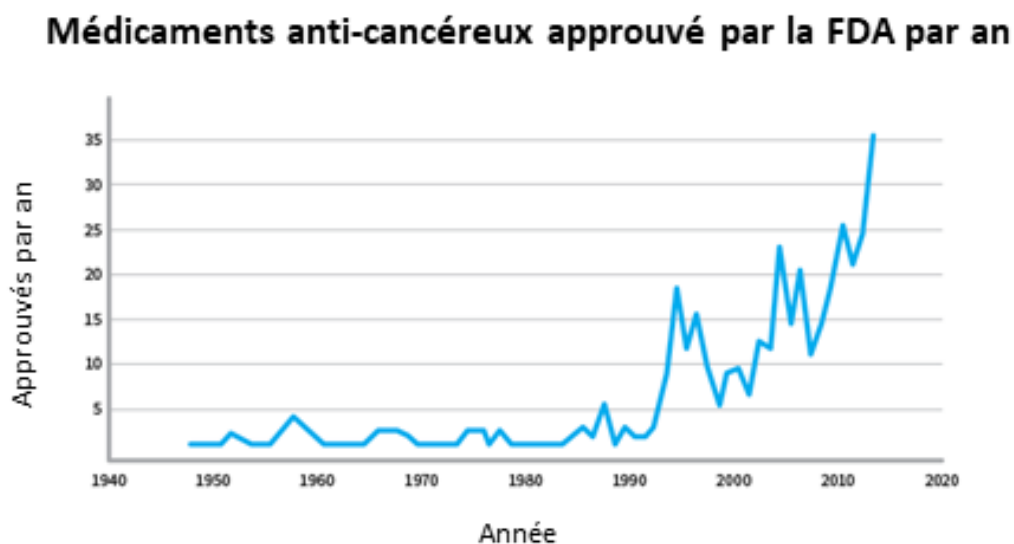


Figure 1-1: Nombre de composés approuvés par la FDA (nouvelles entités moléculaires) au cours du temps⁵.

Le cancer est en fait le nom donné à plusieurs maladies associées et d'origine génétique. Chaque type de cancer est défini par l'organe duquel il provient ou par le type de cellule qui le forme. Au total plus d'une centaine de types de cancer existent⁶. Il a été montré

dans une étude que sur plus d'un million d'échantillons de tumeurs, 2 002 811 de mutations ponctuelles avaient été observées⁷. De plus, la base de données COSMIC (the Catalogue Of Somatic Mutations In Cancer) détaille plus de 6 000 000 de mutations non codantes, 10 534 fusions de gènes et 61 299 réarrangements du génome⁷. Au vu de ces informations, chaque cancer peut presque être considéré comme une maladie unique, c'est pourquoi de plus en plus de traitements ciblés sont mis en place⁸.

En 2000, les chercheurs Hanahan et Weinberg ont caractérisé les cellules cancéreuses au travers de 6 *hallmarks* : l'autosuffisance en signaux de croissance, l'insensibilité aux signaux inhibiteurs de la croissance, la capacité à éviter l'apoptose, la capacité de se répliquer indéfiniment, l'induction de l'angiogenèse et la capacité à former des métastases⁹. En 2011, ceux-ci publient un nouvel article en ajoutant deux nouveaux *hallmarks* : la capacité à éviter une destruction par le système immunitaire et la dérégulation du métabolisme énergétique cellulaire (Figure 1-2)¹⁰.

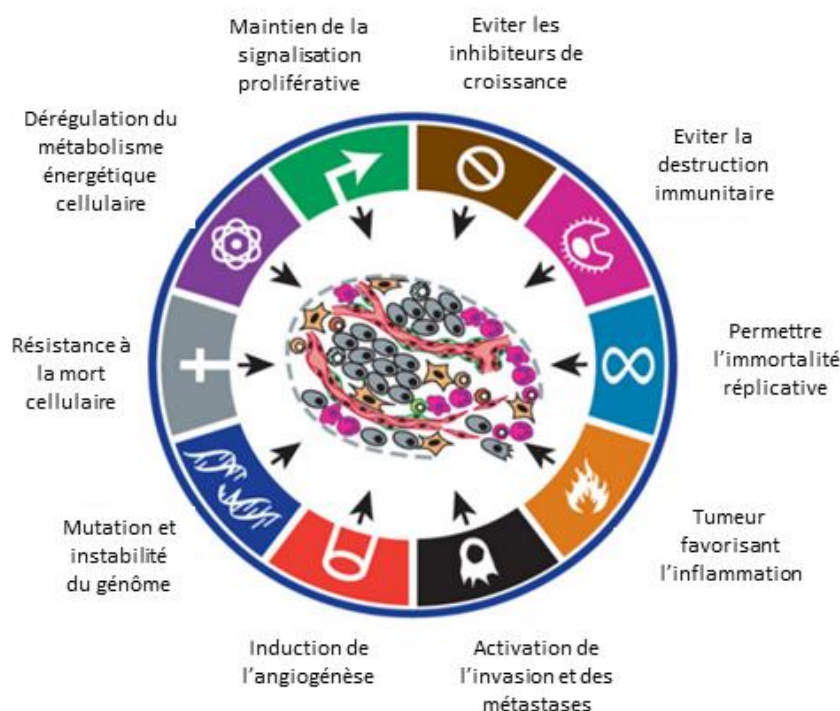


Figure 1-2: Les 8 hallmarks du cancer définis par Hanahan et Weinberg en 2011. Adapté de Hanahan et Weinberg¹⁰.

1.1 La dérégulation du métabolisme énergétique cellulaire

Afin de pouvoir soutenir leur croissance continue et souvent incontrôlée, les cellules cancéreuses doivent ajuster leur métabolisme énergétique. Une cellule saine en condition aérobie va transformer le glucose en pyruvate grâce à la voie de métabolisation de la glycolyse. Celui-ci sera ensuite transformé en CO_2 dans les mitochondries. Ce processus (la phosphorylation oxydative) est le plus efficace pour la production d'ATP. Il permet la production de ~ 18 fois plus d'ATP que la voie de la fermentation lactique, son équivalent en condition anaérobie. En absence d'oxygène, une cellule saine va favoriser cette voie au cours de laquelle le glucose est alors transformé en lactate. Dans les années 1920, Otto Warburg a découvert que les cellules cancéreuses consommaient d'énormes quantités de glucose et que celui-ci était transformé en lactate en condition aérobie^{11,12}. Ce phénomène, tout d'abord nommé la « glycolyse aérobie », est maintenant connu comme « l'effet Warburg » et a permis l'obtention du prix Nobel de physiologie ou médecine pour « sa découverte de la nature et du mode opératoire de l'enzyme respiratoire » à Otto Warburg¹³.

1.1.1 Effet Warburg

Otto Warburg a donc découvert que les cellules cancéreuses consomment une quantité anormalement élevée de glucose et que celui-ci est principalement converti en lactate, et ce indépendamment des concentrations en oxygène. À l'inverse les cellules saines privilégient la phosphorylation oxydative en condition aérobie et la fermentation en condition anaérobie (Figure 1-3). Warburg a attribué la glycolyse aérobie à un dysfonctionnement des mitochondries¹¹. Dans les années qui suivirent, de nombreuses hypothèses furent émises quant à son origine, allant du déséquilibre du pH intracellulaire à la perte d'activité des ATPases. Il fallut attendre de découvrir que la glycolyse aérobie est contrôlée par des facteurs de croissance et que la régulation aberrante de ceux-ci est un événement déclencheur de l'oncogenèse pour écarter certaines de ces hypothèses^{14,15,16-18}. Il a finalement été démontré que l'effet Warburg est nécessaire à la

croissance tumorale^{19,20}. Toutefois la fonction de l'effet Warburg est encore controversée de nos jours²¹.

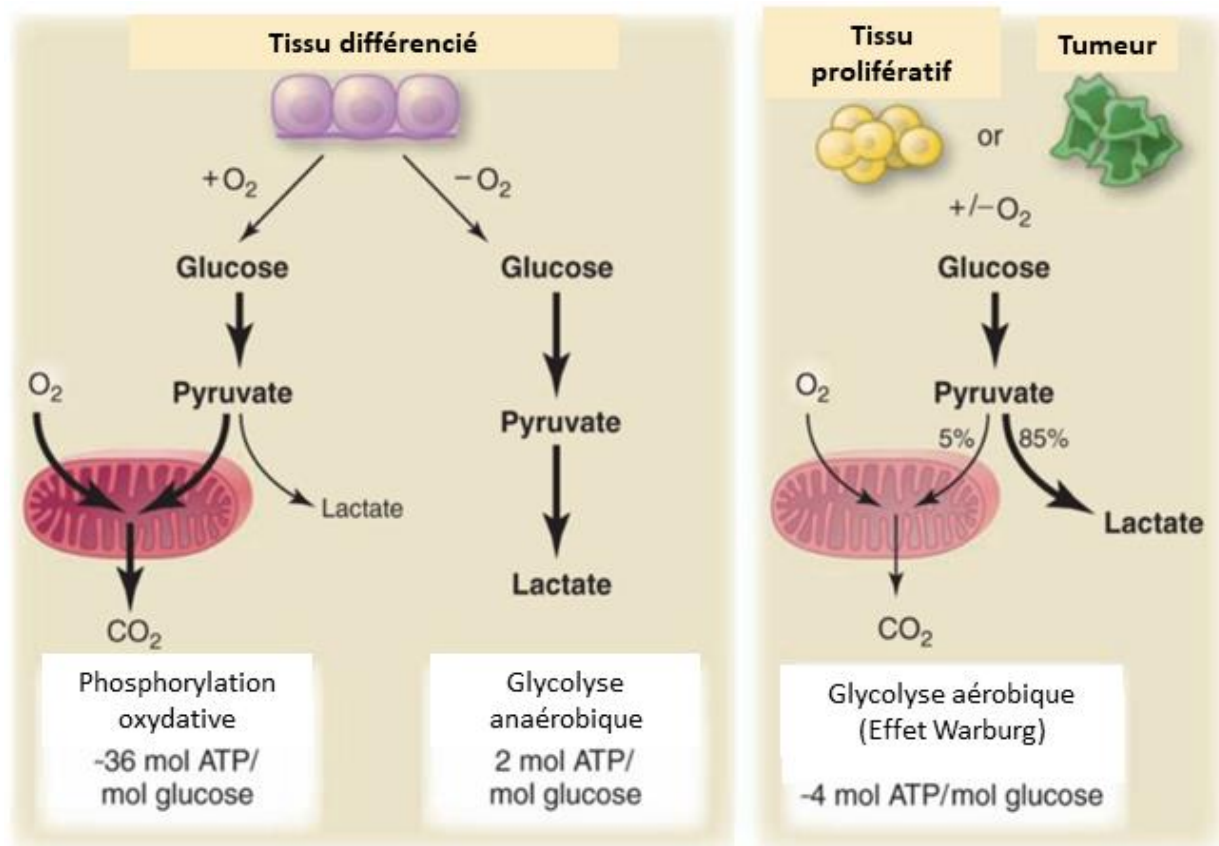


Figure 1-3: Représentation schématique des différences entre la phosphorylation oxydative et la glycolyse anaérobie (dans le cas de cellules saines) et la glycolyse aérobie (dans le cas de cellules cancéreuses). Adapté de Fantin et al.²².

Plusieurs hypothèses ont dès lors été émises quant à la fonction de l'effet Warburg. La première se base sur le fait que bien que la glycolyse aérobie soit moins efficace en termes de production énergétique par rapport à la phosphorylation oxydative, celle-ci est 10 à 100 fois plus rapide. Par conséquent, sur une même durée de temps, la quantité d'ATP produite est comparable quelle que soit la façon dont le glucose est métabolisé²²⁻²⁴. Il a ensuite été montré qu'une cinétique de production d'ATP plus grande pourrait représenter un avantage en cas de compétition pour des ressources limitées, ce qui est observé dans le cadre du microenvironnement tumoral²⁵⁻²⁸.

Une seconde hypothèse porte sur les modifications du microenvironnement tumoral. Une consommation élevée de glucose par la voie de la glycolyse aérobie diminue le pH du microenvironnement suite à la sécrétion de lactate (Veronica Estrella, Tingan Chen, Mark Lloyd 2013). Il a été montré qu'un microenvironnement acide altère l'interface tumeur-stroma, ce qui permet d'améliorer l'invasion par les cellules cancéreuses (Veronica Estrella, Tingan Chen, Mark Lloyd 2013; Robert A. Gatenby 1996). Enfin, l'utilisation de cette voie induit une compétition entre les tumeurs³¹ et les lymphocytes infiltrateurs de tumeurs pour le glucose, dont la disponibilité est limitée^{27,28}. Bien qu'il soit probable que l'effet Warburg apporte un avantage aux cellules cancéreuses du point de vue du microenvironnement, celui-ci est supposé arriver tôt durant le phénomène d'oncogenèse. De plus il a été montré que dans des conditions isolées, la glycolyse aérobie reste la principale voie de dégradation du glucose³². Il est donc peu probable que le bénéfice conféré au niveau de l'environnement tumoral soit la raison principale de l'utilisation de cette voie en particulier.

Une troisième hypothèse propose que l'effet Warburg confère des fonctions de signalisation directe aux cellules tumorales. Il peut causer des modifications dans le potentiel redox des mitochondries modifiant la génération de *reactive oxygen species*²³. Le NADH est aussi un acteur important du potentiel redox dans les mitochondries grâce à sa capacité de transporteur d'électron. Toutefois dans le cas d'une glycolyse plus rapide, les moyens mis en place pour maintenir l'équilibre du NADH (malate aspartate *shuttle*) ne sont plus capables de compenser la génération de NAD⁺ entraînée par la conversion du pyruvate en lactate³³. Cette conséquence de l'effet Warburg peut aussi affecter la production d'*oncogene-induced senescence* (OIS)³⁴. Les OIS, qui possèdent des fonctions suppressives de tumeur, sont en effet régulées par l'oxydation du glucose.

Finalement, une dernière hypothèse stipule que la consommation supérieure en glucose est utilisée comme source de carbone dans les voies anaboliques supportant la prolifération cellulaire^{22,35-38}. Un premier argument est que les cellules cancéreuses plutôt que d'avoir une demande limitante en ATP, ont besoin de son équivalent réduit, le NADPH. Une plus grande consommation de glucose permet une synthèse en plus grande quantité du NADPH. Un second argument est que la glyceraldéhyde

déhydrogénase, qui maintient la glycolyse active, consomme le NAD^+ en le transformant en NADH. La glycolyse aérobie permet de régénérer le NAD^+ lors de la transformation du pyruvate en lactate. Finalement nous pouvons aussi mettre en avant les intermédiaires glycolytiques qui sont déviés vers des voies secondaires comme la voie de la sérine²³. Le flux glycolytique est en partie détourné vers la biosynthèse *de novo* de la sérine par l'enzyme phosphoglycérate déhydrogénase.

1.2 La voie de la sérine

1.2.1 Généralités

La sérine est un acide aminé non essentiel, c'est-à-dire que celui-ci peut être obtenu dans notre alimentation. Toutefois, la synthèse *de novo* est rendue possible par la voie de biosynthèse de la sérine au sein des cellules. Celle-ci détourne le 3-phosphoglycérate (3-PG), un intermédiaire glycolytique, de la voie de la glycolyse (Figure 1-4). La première enzyme de la voie, la phosphoglycérate déhydrogénase (PHGDH), va oxyder le 3-PG en 3-phosphohydroxypyruvate (PPyr) en présence de NAD^+ qui sera réduit en NADH. Cette réaction est suivie par la transamination du PPyr pour former la phosphosérine (Pser) qui est réalisée par la phosphosérine aminotransférase (PSAT1). Le glutamate va donner son groupement amine et la molécule formée, l' α -cétoglutarate, va être utilisé par le cycle TCA. La sérine va finalement être formée après déphosphorylation par la phosphosérine phosphatase. PHGDH est souvent considérée comme l'enzyme limitante de cette voie^{39,40}.

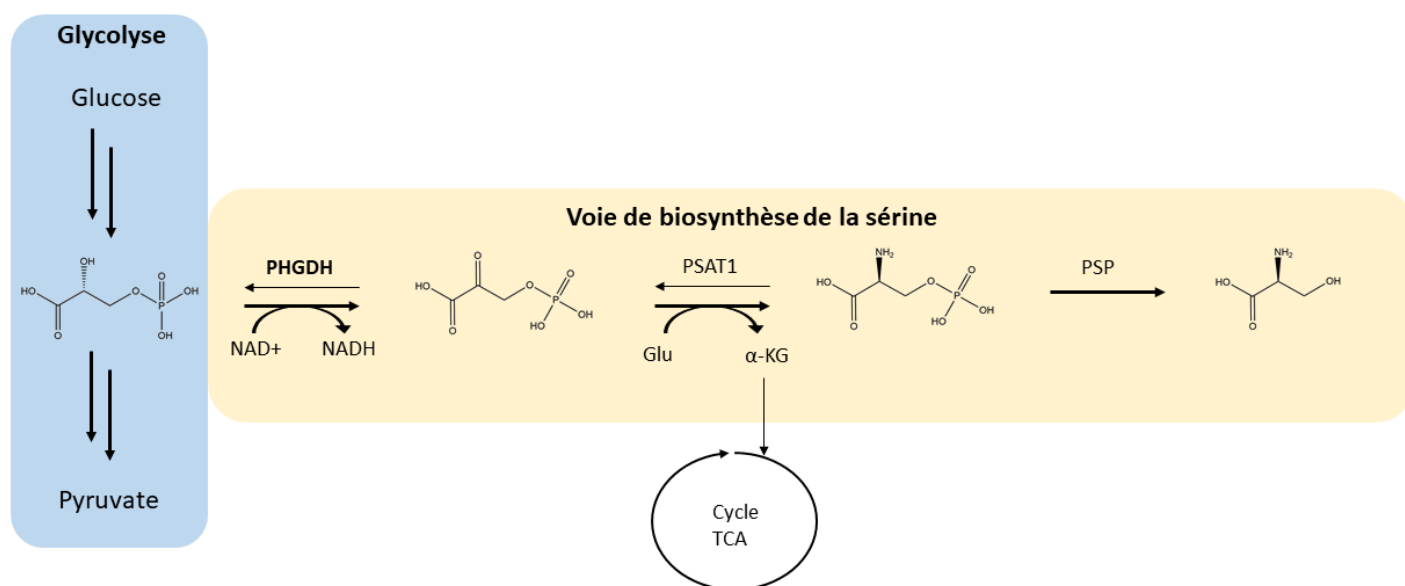


Figure 1-4: Voie de biosynthèse de la sérine. PHGDH = phosphoglycérate déshydrogénase ; PSAT1 = phosphosérine aminotransférase ; PSP= Phosphosérine phosphatase ; Glu = glutamate ; α -KG = α -cétoglutarate.

1.2.2 Biosynthèse de la sérine par les cellules saines

Bien que la sérine soit un acide aminé non essentiel et est par conséquent retrouvé dans le plasma, la synthèse *de novo* de celle-ci contribue pour une part non négligeable à la sérine disponible *in vivo*^{41,42}. La sérine présente dans le plasma présente un faible transport au travers de la barrière hématoencéphalique et la quantité de sérine provenant de notre régime alimentaire peut s'avérer insuffisante pour supporter les fonctions cérébrales⁴³. Par conséquent, la synthèse *de novo* de la sérine dans cette région est nécessaire. Il a été montré que des déficiences en enzyme PHGDH, ou un épuisement de la sérine lié à des mutations dans sa voie de biosynthèse peuvent entraîner des désordres neurologiques. Ceux-ci peuvent être traités en augmentant la quantité de sérine présente dans le régime alimentaire⁴⁴⁻⁴⁶.

La biosynthèse de la sérine est une voie fortement régulée. Cependant, elle peut être induite par différents stress. Un épuisement en sérine ou une accumulation en 2-phosphoglycérate (le produit de réaction du 3-PG dans la voie de la glycolyse) permettent l'activation de la voie⁴⁷.

1.2.3 Implication de la sérine dans les voies métaboliques

La sérine, et donc sa voie de biosynthèse, sont impliquées dans plusieurs voies de synthèse (Figure 1-5)(inspiré de la revue de N. Hay⁴⁸).

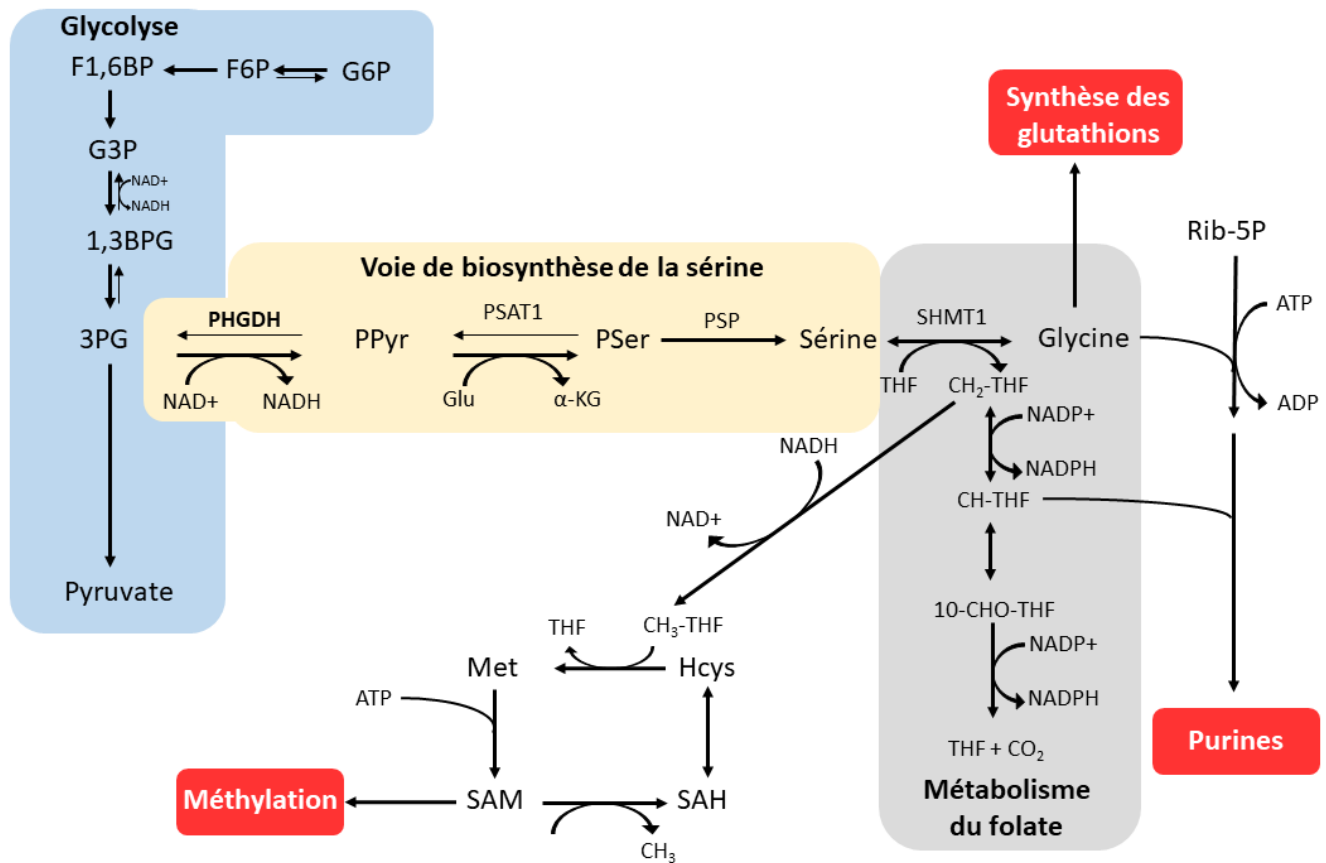


Figure 1-5: Représentation de la voie de biosynthèse de la sérine et son implication dans le métabolisme à un carbone, le cycle de la méthionine, la voie de biosynthèse de la purine et la génération de glutathion. PHGDH = phosphoglycérate déshydrogénase ; PSAT1 = phosphosérine aminotransférase ; PSP = Phosphosérine phosphatase ; Glu = glutamate ; α-KG = α-cétoglutarate ; GSH = glutathion réduit ; THF = tétrahydrofolate ; SHMT1 = sérinehydroxyméthyltransférase. Adapté de Hay et al.⁴⁸.

La sérine peut tout d'abord être convertie en glycine par les sérinehydroxyméthyltransférases afin de participer au métabolisme à un carbone. Cette réaction s'accompagne du transfert d'un groupement méthylène au tétrahydrofolate (THF). La glycine est impliquée dans la synthèse directe des glutathions participant dès lors à la régulation du potentiel rédox au sein du cytosol.

Le THF peut être converti en 5-méthyl-THF (CH₂-THF) afin de fournir le groupement méthyle à la S-adénosyl méthionine (SAM). Cette molécule est responsable de la méthylation de l'ADN, de métabolites et d'une série de protéines.

Le CH₂-THF peut aussi être transformé en 5,10-méthényl-THF (CH-THF) par l'enzyme méthylènetétrahydrofolate déshydrogénase (MTHFD1) dans le cytoplasme ou MTHFD2 dans la mitochondrie afin d'être utilisé dans le métabolisme du folate. Il sera ensuite converti en 10-formyl-THF (10-CHO-THF). Il a été montré que le métabolisme du folate est une source importante de NADPH pour les cellules cancéreuses^{49,50}.

Finalement le CH-THF formé par les MTHFD peut être intégré dans la formation de purines et par conséquent montre l'implication de la sérine/glycine dans la synthèse d'ADN, ARN, ATP, et bien d'autres métabolites dont les purines sont à l'origine.

Il a été montré que lors de l'inhibition de PHGDH (première enzyme de la voie de la sérine), la glycine est convertie en sérine par la SHMT1⁵¹. Cette réaction est aussi observée dans le cas de glycine dérivée de sérine exogène. Le CH₂-THF étant consommé, la synthèse de purine est réduite⁵¹. Cette observation pourrait expliquer pourquoi la synthèse *de novo* de sérine est nécessaire à la prolifération cellulaire.

1.2.4 La voie de la sérine dans les mécanismes cancéreux

En 2011, deux études indépendantes ont suscité un intérêt thérapeutique nouveau pour la voie de la sérine^{52,53}. Celles-ci ont montré que PHGDH était surexprimée dans une série de mélanomes et cancers du sein (principalement dans les cancers ER-négatifs). Possemato et ses collègues ont montré que les cellules surexprimant PHGDH présentent un flux augmenté vers la voie de la sérine par rapport aux cellules ne surexprimant pas la première enzyme de la voie. Lors de la privation de sérine, il a été montré que les cellules surexprimant PHGDH conservent une prolifération robuste alors que pour les cellules présentant un niveau faible de PHGDH, la prolifération est fortement ralentie voire complètement stoppée. Des cellules saines ne surexprimant naturellement pas PHGDH (MCF-10A) présentent une prolifération amplifiée en absence de sérine après

avoir été modifiées afin d'acquérir la capacité de produire un niveau en PHGDH similaire aux cellules cancéreuses la surexprimant. Cette expérimentation a permis de montrer que la surexpression de PHGDH est suffisante pour rediriger le 3-PG vers la voie de biosynthèse de la sérine et par conséquent supporter la prolifération cellulaire. Une expérimentation similaire a été effectuée dans le cas de PSP mais la surexpression de cette enzyme n'est pas suffisante pour rediriger le flux et soutenir la prolifération. Cette dernière expérimentation met en évidence que PHGDH est la cible privilégiée afin d'empêcher la synthèse *de novo* de sérine dans les cellules cancéreuses. Cette étude conclut que 70% des cancers du sein ER-négatifs testés présentent une expression élevée en PHGDH et pourraient être sensibles à un inhibiteur de PHGDH⁵³. Il a été mis en évidence que des cellules présentant de hauts niveaux en PHGDH sont incapables de proliférer dans des milieux contenant un niveau normal de sérine si leur production de PHGDH a été inhibée. L'hypothèse la plus probable pour les auteurs est que les enzymes PHGDH, PSAT1 et PSP produisent des métabolites, en plus de la sérine, qui sont eux aussi essentiels pour la prolifération cellulaire. Cette observation rend encore plus attractif le ciblage de cette voie car les effets secondaires des inhibiteurs pourraient être compensés par un régime enrichi en sérine.

1.3 La phosphoglycérate déhydrogénase (PHGDH)

Dans la section précédente, nous avons montré l'intérêt que présente PHGDH humaine en tant que cible thérapeutique potentielle dans le traitement du cancer. Nous décrivons ici les informations connues pour cette enzyme. Celles-ci étant encore partielles, nous les compléterons avec des informations provenant d'autres souches semblables à la forme humaine de PHGDH.

1.3.1 Informations générales sur PHGDH

La phosphoglycérate déhydrogénase est la première enzyme de la voie de biosynthèse de la sérine. Elle catalyse l'oxydation du 3-phosphoglycérate en phosphohydroxypyruvate en présence de NAD⁺ qui est réduit en NADH. L'enzyme est capable de réaliser la réaction dans le sens de l'oxydation et de la réduction du substrat. L'enzyme catalyse préférentiellement la réaction de réduction du phosphohydroxypyruvate en 3-phosphohydroxypyruvate. A l'équilibre, moins de 5% du substrat/produit se trouve sous la forme de phosphohydroxypyruvate⁵⁴. Dans la cellule, la présence de l'enzyme PSP, qui catalyse une réaction complète par déphosphorylation, permet de favoriser la production de phosphohydroxypyruvate. Finalement, il a été supposé sur base de résultats de cristallisation et de *Differential Scanning Fluorimetry* (DSF) que l'ordre de liaison au sein du site actif imposerait la liaison du cofacteur avant celle du substrat⁵⁵.

L'enzyme est constituée de 533 acides aminés et présente une masse de 56,7Da⁵⁶.

Les enzymes PHGDH provenant de différentes souches sont classées selon trois catégories. Les différents types de PHGDH sont déterminés en fonction de leur structure (Figure 1-6)⁵⁴.

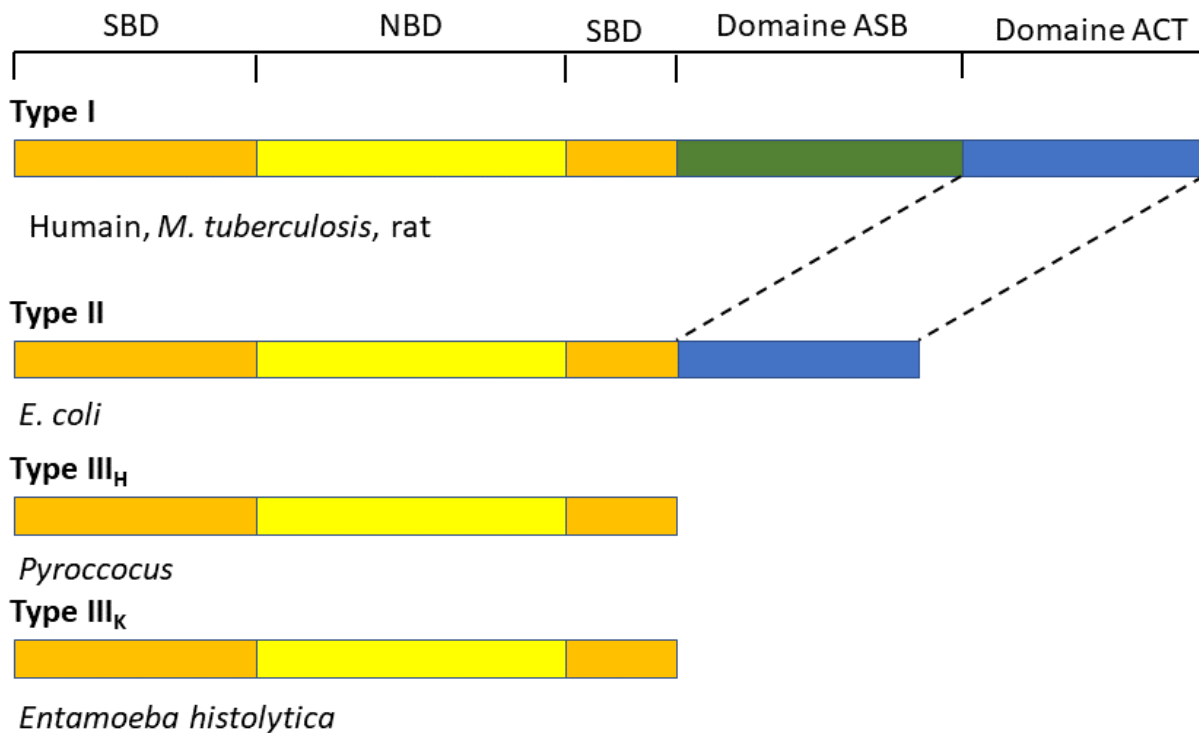


Figure 1-6: Classification des PHGDH provenant de différents organismes. La liste des différents organismes n'est pas exhaustive. Adapté de Grant et al.⁵⁴.

Toutes les PHGDH existantes présentent un domaine de liaison au substrat (*Substrat Binding Domain* - SBD) et un domaine de liaison au cofacteur (*Nucleotide Binding Domain* - NBD). Les enzymes présentes chez les organismes les moins évolués ne possèdent que ces deux domaines et sont classés de Type III. Les sous-classifications H et K correspondent à la modification d'une histidine (H) en une lysine (K) dans le site actif.

Les enzymes de Type II présentent un domaine supplémentaire, le domaine ACT. Ce motif récurrent qui a été identifié en 1999, a été nommé selon la première lettre des trois premières enzymes dans lesquelles il a été identifié (Aspartate kinase, Chorismate mutase et TyrA)⁵⁷. Ce domaine de 60-70 acides aminés présente généralement une fonction régulatrice qui est activée par la liaison d'un acide aminé. Celui-ci semble fonctionnel uniquement chez les bactéries. Ce domaine présente généralement des mutations au niveau des acides aminés critiques quand il est présent sur des enzymes provenant de mammifère. C'est le cas pour la forme humaine de PHGDH.

Finalement, les PHGDHs de Type I sont présentes dans les organismes les plus évolués. Celles-ci sont caractérisées par la présence d'un domaine ASB en plus des domaines précédemment cités. Ce domaine de 150 acides aminés a été observé pour la première fois sur la PHGDH provenant de la bactérie *Mycobacterium tuberculosis* (tbPHGDH)⁵⁸. Peu d'informations sont actuellement disponibles sur ce domaine.

La forme humaine de PHGDH est donc de type I et présente les quatre domaines spécifiques aux enzymes de type I : NBD, SBD, ACT et ASB.

1.3.2 Structure de l'enzyme

La structure de l'enzyme complète de PHGDH humaine n'a, à ce jour, pas encore été résolue. Des formes tronquées de l'enzyme ont toutefois permis d'apporter des informations structurales de celle-ci. Deux formes tronquées de l'enzyme humaine existent. La première que nous nommerons *short* PHGDH (sPHGDH) comporte les acides aminés 3 -314 qui correspondent aux domaines SBD et NBD. Quatre structures cristallographiques de cette forme sont présentes sur la *Protein Data Bank* (PDB). Deux structures présentent un ligand dans le site de liaison du substrat : D-malate (2G76) et L-tartrate (5N6C). Les deux autres structures (6PLG, 6PLF) qui ont été publiées durant la rédaction de cette thèse présentent, chacune un inhibiteur dans la poche de liaison du cofacteur (discuté à la section 1.3.3)⁵⁹.

La structure de sPHGDH présente dans son unité asymétrique un homodimère (Figure 1-7). Unterlass *et al.* ont pu cristalliser sPHGDH en présence de tartrate et obtenir deux monomères structuralement différents au sein de l'unité asymétrique (Figure 1-7)⁵⁵. Le premier, lié au L-tartrate, présente une rotation de 29° au niveau du domaine de liaison au substrat par rapport au second qui n'a pas lié le L-tartrate. Le mode de liaison du tartrate au sein du site actif a été mise en évidence. Celui-ci forme des interactions avec R235, H282 du domaine fixe et R53, S54, R74, T77 du domaine flexible. Un mode d'interaction similaire est observé pour la liaison du malate. Ces deux structures contiennent le cofacteur NAD⁺.

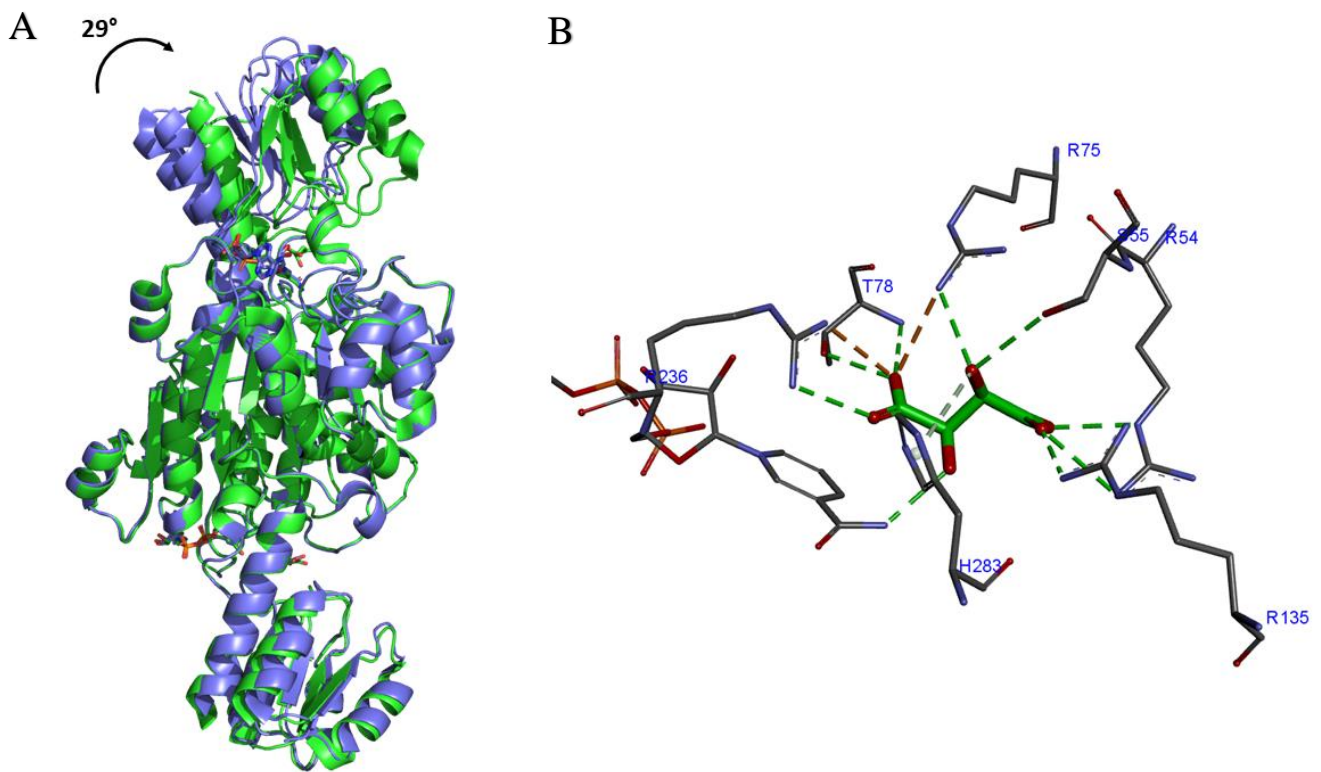


Figure 1-7: Structure cristallographie de sPHGDH co-cristallisée avec le L-tartrate (5N6C). (A) Alignement des deux monomères présents dans l'unité asymétrique, sPHGDH en présence de L-tartrate (mauve) et en absence de L-tartrate (gris). (B) Mode de liaison du L-tartrate au sein de la poche. Adapté de Unterlass et al.⁵⁵.

La seconde forme tronquée est composée des acides aminés 93-298, éliminant par conséquent la majeure partie du domaine de liaison au substrat⁶⁰. Les sept structures qui ont été cristallisées avec cette enzyme présentent un fragment lié au sein de la poche de liaison au substrat. Contrairement à sPHGDH, l'unité asymétrique contient un monomère de l'enzyme.

La forme humaine de PHGDH n'est pas la seule à avoir été cristallisée. Le Tableau 1-1 reprend toutes les formes de PHGDH ayant été cristallisées et déposées sur la PDB.

Tableau 1-1: Liste des structures de PHGDH présentes sur la PDB. Pour chaque structure il est indiqué l'organisme, le type, la spécificité de la structure et le code PDB associé à celle-ci.

Organisme	Type	Ligand ou mutation	Code PDB
Humain (3 - 314)	I	NAD+ / Malate	2G76
Humain (3 - 314)	I	NAD+ / Tartrate	5N6C ⁵⁵
Humain (3 - 314)	I	Composé 1 ^a	6PLF ⁵⁹
Humain (3 - 314)	I	Composé 15 ¹	6PLG ⁵⁹
Humain (93 - 298)	I	N-(3-chloro-4-methoxyphenyl) acétamide	5N53 ⁶⁰
Humain (93 - 298)	I	1-methyl-3-phenyl-1H-pyrazol-5-amine	5NZO ⁶⁰
Humain (93 - 298)	I	3-Hydroxybenzisoxazole	5NZP ⁶⁰
Humain (93 - 298)	I	3-(1,3-oxazol-5-yl)aniline.	5NZQ ⁶⁰
Humain (93 - 298)	I	5-fluoro-2-methylbenzoic acide	5OFV ⁶⁰
Humain (93 - 298)	I	3-Chloro-4-fluorobenzamide	5OFW ⁶⁰
Humain (93 - 298)	I	5-amino-1-methyl-1H-indole	5OFM ⁶⁰
<i>M. tuberculosis</i>	I	Serine	3DC2 ⁶¹
<i>M. tuberculosis</i>	I	Phosphohydroxypyruvate	3DDN ⁶¹
<i>M. tuberculosis</i>	I	Tartrate	1YGY ⁵⁸
<i>E. coli</i>	II	Serine / G336V	2P9C ⁶²
<i>E. coli</i>	II	G336V	2P9E ⁶²
<i>E. coli</i>	II	G336V, G337V	2P9G ⁶²
<i>E. coli</i>	II	Serine / G336V	2PA3 ⁶²
<i>E. coli</i>	II	NAD+, 2-oxoglutarate	1YBA ⁶³
<i>E. coli</i>	II	NAD+	1SC6 ⁶⁴
<i>E. coli</i>	II	NAD+, serine	1PSD ⁶⁵
<i>Pyrococcus horikoshii</i>	III _H	NAD+	1WWK
<i>Entamoeba histolytica</i>	III _K		4NFY ⁶⁶
<i>Entamoeba histolytica</i>	III _K	Phosphoglycérate	4NJM ⁶⁶
<i>Entamoeba histolytica</i>	III _k	NAD+	4NJO ⁶⁶
<i>Ralstonia solanacearum</i>			3GG9
<i>Lactobacillus plantarum</i>			3EVT

^a La structure du composé chef de file de la série est présenté à la Figure 1-10

Le Tableau 1-1 montre que l'enzyme humaine est actuellement la plus étudiée et que si la forme humaine est intéressante d'un point de vue thérapeutique, les autres organismes ont été étudiés par une approche fonctionnelle. Parmi les mammifères, seule la forme humaine de PHGDH a déjà été cristallisée. Parmi les enzymes de Type I, seule celle de *Mycobacterium tuberculosis* a pu être cristallisée sous forme complète.

Les formes humaine et bactérienne (tbPHGDH) sont semblables. En effet, 32% d'identité de séquence sont observés entre les deux enzymes^b. La forme bactérienne, ainsi que son homonyme humain, sont actives sous forme tétramérique. La structure cristallographie de tbPHGDH sous forme d'un tétramère asymétrique est présentée à la Figure 1-8. Grant et *al.* ont mis en évidence que l'interaction observée entre les deux domaines ACT (Figure 1-8B, R-R) est responsable de la formation du tétramère⁵⁸. Il est probable que les caractéristiques retrouvées chez tbPHGDH soient partagées avec les autres enzymes PHGDH de Type I.

^b Déterminé par alignement de séquence sur le site <https://www.uniprot.org>. Programme utilisé : CLUSTALO

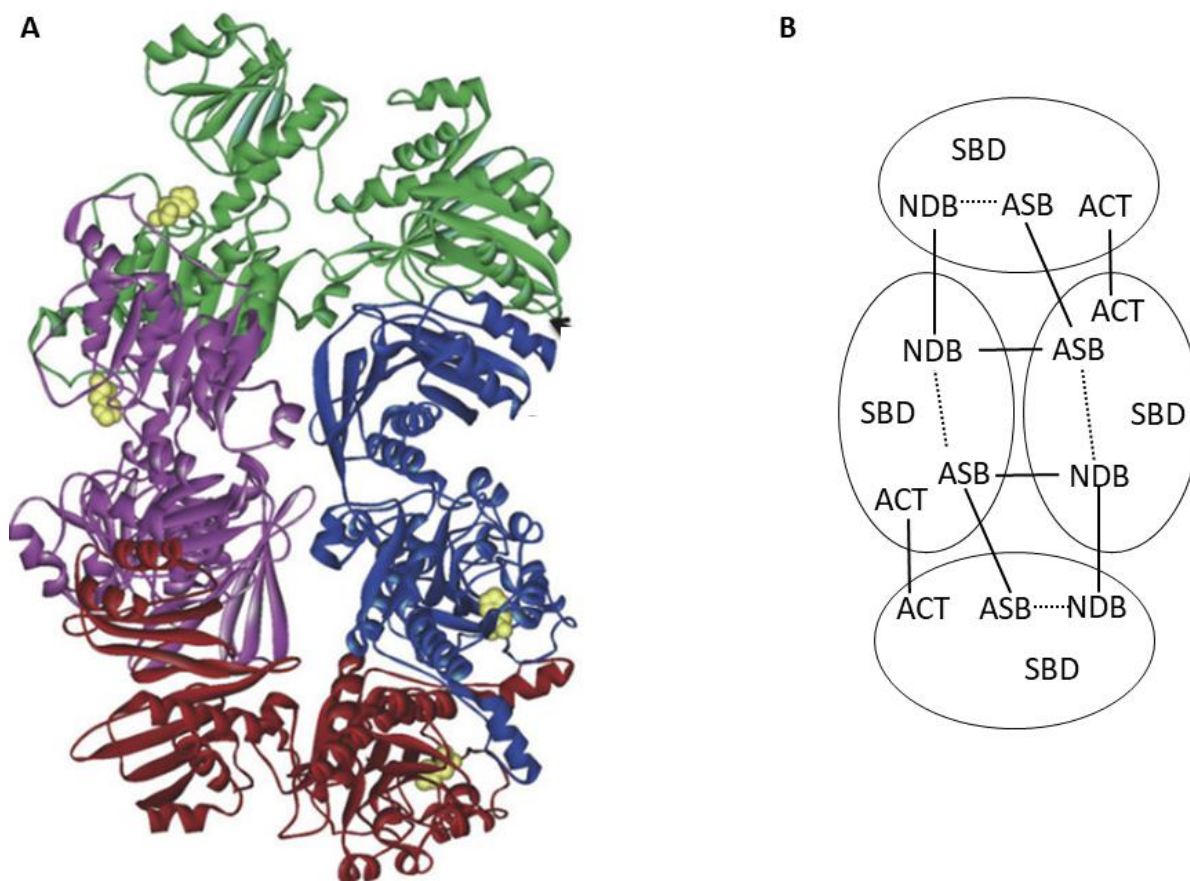


Figure 1-8: Structure de *tbPHGDH* (3DDN). (A) Tétramère de *tbPHGDH* (chaque tétramère est représenté dans une couleur différente, la position du site actif est représentée par les balles jaunes). (B) Représentation schématique des différents domaines et de leur position relative au sein du tétramère. Adapté de Grand et al.^{54,58}.

1.3.3 Inhibiteurs ciblant PHGDH

Comme précédemment mentionné, PHGDH a été identifiée en 2011 comme une cible thérapeutique potentielle. Depuis lors, plusieurs inhibiteurs ciblant cette enzyme ont été conçus. Après avoir réalisé le criblage de plus de 800.000 composés, le premier inhibiteur de PHGDH, le CBR-5884, a été identifié et publié en 2016 (Figure 1-9)⁶⁷. Ce composé est un inhibiteur non compétitif capable de perturber la forme tétramérique de l'enzyme (forme active de l'enzyme) en favorisant la formation d'un dimère. L'inhibiteur présente un $IC_{50} = 33 \pm 12 \mu M$. Le NCT-503 (Figure 1-9) a ensuite été publié et présente un $IC_{50} = 2,5 \pm 0,6 \mu M$ ⁵¹. Il a aussi été identifié suite au criblage de plus de 400.000 composés. Tout comme le CBR-5884, son mode de liaison avec l'enzyme n'est pas connu et il est non compétitif du substrat et du cofacteur. Ces deux composés

présentent une activité *in vivo* et nous observons en leur présence une diminution de la quantité de sérine et glycine intracellulaires. Des dérivés α -cétothioamides (Figure 1-9) présentent également une inhibition sur l'enzyme. Le principal composé de cette série présente un $IC_{50} = 30,9\mu M^{68}$. Il s'agit d'un inhibiteur non compétitif. Sa liaison entraîne également une inhibition de la prolifération des cellules cancéreuses quand il est mis en leur présence. Il est intéressant de noter que le mode de liaison à l'enzyme n'est connu pour aucun de ces composés.

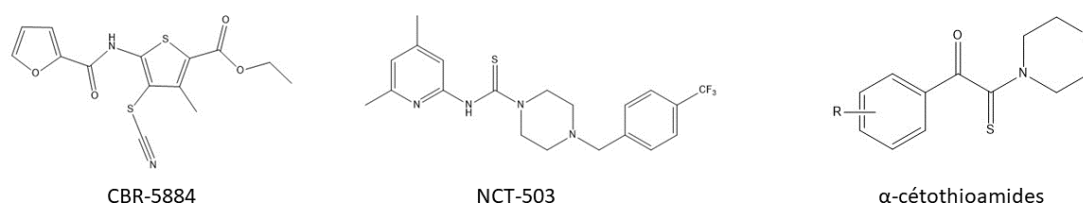
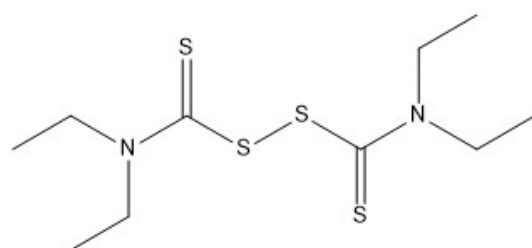
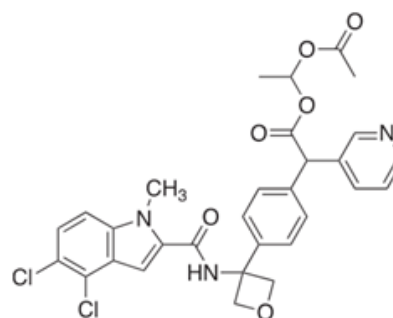


Figure 1-9: Structure des premiers inhibiteurs de PHGDH : CBR-5884, NCT-503, dérivés α -cétothioamides.

Plusieurs autres composés comportant des informations quant à leur liaison avec l'enzyme ont ensuite été publiés. Nous avons pu mettre en évidence avec nos collaborateurs que le disulfirame (Figure 1-10) interfère, tout comme le CBR-5884, avec la formation de la forme tétramérique de l'enzyme et qu'il forme un lien covalent entre son adduit (rupture de la liaison S-S) et la C116⁶⁹. Un $IC_{50} = 0,59\mu M$ a été mesuré pour ce composé. Très récemment, les seconds inhibiteurs compétitifs du site de liaison du NAD⁺ pour lesquels le mode d'action a été identifié ont été publiés par Mullarky *et al.*⁵⁹. Ces composés sont inspirés du premier inhibiteur compétitif du site de liaison au NAD⁺ découvert par Astrazeneca mais pour lesquelles les structures cristallographiques n'ont jamais été publiées⁷⁰. Comme décrit dans la section 1.3.2 (p.15), ces auteurs ont réussi à obtenir une structure cristallographique avec la forme tronquée de l'enzyme pour deux inhibiteurs. Le composé **18**, qui est le plus prometteur, présente un IC_{50} de $1,46\mu M$ et a été co-cristallisé dans la poche de liaison au cofacteur. Ce composé peut être métabolisé par hydrolyse de sa fonction cétone afin de former la forme la plus active de la série (composé **15**).



Disulfirame



Composé 18

Figure 1-10: Structure des premiers inhibiteurs de PHGDH dont l'interaction avec l'enzyme est connue : disulfirame, composé 18.

Le disulfirame et le composé **18**, deux inhibiteurs publiés au courant de l'année 2019, montrent que des informations structurales et fonctionnelles de l'enzyme PHGDH sont nécessaires afin d'avoir une meilleure compréhension du mode de liaison des inhibiteurs connus et la possibilité de concevoir de nouveaux inhibiteurs sélectifs du site actif de PHGDH.

2. Objectif et stratégie

2 Objectif et stratégie

La phosphoglycérate déshydrogénase humaine est une cible thérapeutique intéressante. Ces dernières années, un nombre grandissant de recherches a mis en évidence l'implication de la voie de la sérine et de l'enzyme PHGDH dans les mécanismes cancéreux. Par conséquent, un nombre grandissant de publications visant à l'inhibition de cette enzyme ont vu le jour. Toutefois, le manque important d'informations concernant la structure et le fonctionnement de cette enzyme représente toujours un obstacle au développement d'inhibiteurs efficaces.

Afin de combler ce manque de connaissances, ce travail a pour but d'apporter des informations structurales et fonctionnelles complémentaires sur PHGDH afin de permettre de mieux comprendre comment l'inhiber, ou améliorer l'inhibition des inhibiteurs connus de cette enzyme. Notre travail ciblera principalement une meilleure compréhension de l'enzyme sous sa forme active (tétramérique) et de son site actif comprenant la liaison au cofacteur et à son substrat.

Nous commencerons par l'approche la plus courante dans le cadre de la résolution de la structure d'enzyme : la cristallographie. L'étape limitante de la cristallographie est généralement l'obtention de cristaux diffractant suffisamment. Plusieurs méthodes sont envisagées pour favoriser la production de cristaux. En plus de modifications standards pouvant être apportées à la protéine (protocole d'expression et de purification, paramètres de cristallisation, etc.), nous envisageons l'ajout d'un *Maltose Binding Protein* (MBP). Cette protéine de fusion, qui présente l'avantage d'améliorer la solubilité de la protéine, peut aussi induire la formation de cristaux. Des modifications en surface de la protéine (réduction d'entropie de surface, SER) peuvent favoriser les contacts interprotéines et entraîner la cristallisation d'une enzyme récalcitrante. Si aucune de ces méthodes ne donnait de résultats, nous travaillerions avec la protéine tronquée (sPHGDH) pour obtenir des informations supplémentaires sur le site actif.

Afin de pouvoir étudier l'enzyme sous forme tétramérique, nous pouvons générer un modèle se basant sur les homologues bactériens déjà cristallisés. Si cette méthode nous

permet d'obtenir un modèle plausible, elle pourrait être combinée à de la spectroscopie de masse couplée au cross-linking afin de confirmer les interactions observées dans notre modèle.

Nous tenterons ensuite d'obtenir des informations sur la liaison du substrat à l'aide de la cristallographie de l'enzyme complète ou tronquée ou bien du modèle généré en fonction des résultats obtenus. La combinaison de cette approche avec l'induction de mutations au sein du site actif permettra de mieux comprendre le rôle des acides aminés impliqués dans la liaison du substrat ou de l'activité de l'enzyme.

Finalement, une approche par *docking* sera réalisée afin de trouver de nouveaux ligands pouvant se lier dans les deux poches de liaison du site actif et mettre en évidence les interactions privilégiées au sein de ceux-ci.

Au terme de cette étude, nous espérons pouvoir présenter des informations structurales nouvelles concernant la forme tétramérique de PHGDH. Et des informations fonctionnelles sur les deux poches de liaison du site actif.

3. Expression, purification et cristallisation de PHGDH

3 Expression, purification et cristallisation de PHGDH

La méthode de choix pour la résolution de la structure est la cristallographie. Toutefois, la production de cristaux de protéine est l'étape limitante pour l'obtention de sa structure tridimensionnelle. En effet, celle-ci demande une grande pureté (> 98%), de grandes quantités d'enzyme et un échantillon homogène. Ce chapitre revient sur la mise en place du protocole d'expression et de purification de l'enzyme, sur la stratégie de cristallisation et enfin sur les moyens mis en place pour aider à la cristallisation de celle-ci.

3.1 PHGDH sous sa forme sauvage

Afin d'augmenter nos chances d'obtenir des cristaux de PHGDH, plusieurs protocoles d'expression et de purification ont été testés et mis au point. Le protocole décrit ci-dessous est celui qui a permis l'obtention du cristal qui a diffracté.

Le plasmide permettant l'expression de PHGDH nous a été envoyé par le Dr. Edouard Mullarky de l'université de Cornell (Etats-Unis). Ce plasmide est composé d'un vecteur pET-28a et du gène codant pour la forme sauvage de l'enzyme (Figure 3-1). Le vecteur est constitué d'un gène de résistance à la kanamycine permettant de sélectionner les colonies transformées, d'une origine de réplication, d'un site T7 promoteur et terminateur afin de contrôler l'expression de l'enzyme et d'un site ajoutant un His₆-tag sur la partie N-terminale de l'enzyme.

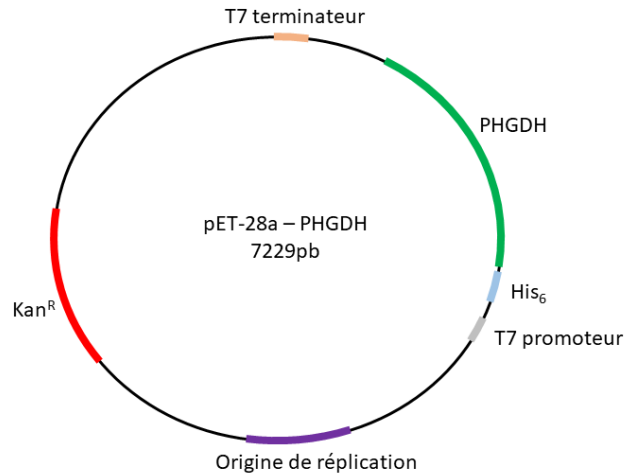


Figure 3-1: Schéma représentant le plasmide constitué du vecteur pET-28a et le gène codant pour PHGDH sous sa forme complète.

Le plasmide a été inséré dans des bactéries *E. coli* BL21(DE3) thermo-compétentes. La kanamycine a permis de sélectionner les colonies transformées. Celles-ci ont été mises en culture et l'expression de la protéine a été induite par ajout d'IPTG. Une lyse par sonication est effectuée afin d'éclater les parois bactériennes. La protéine d'intérêt solubilisée dans le surnageant de lyse est purifiée par chromatographies successives. La première est une chromatographie d'affinité qui va lier l'étiquette histidine présente sur la protéine à l'aide du nickel fixé sur la colonne (HisTrap FF crude, GE healthcare) (Figure 3-2).

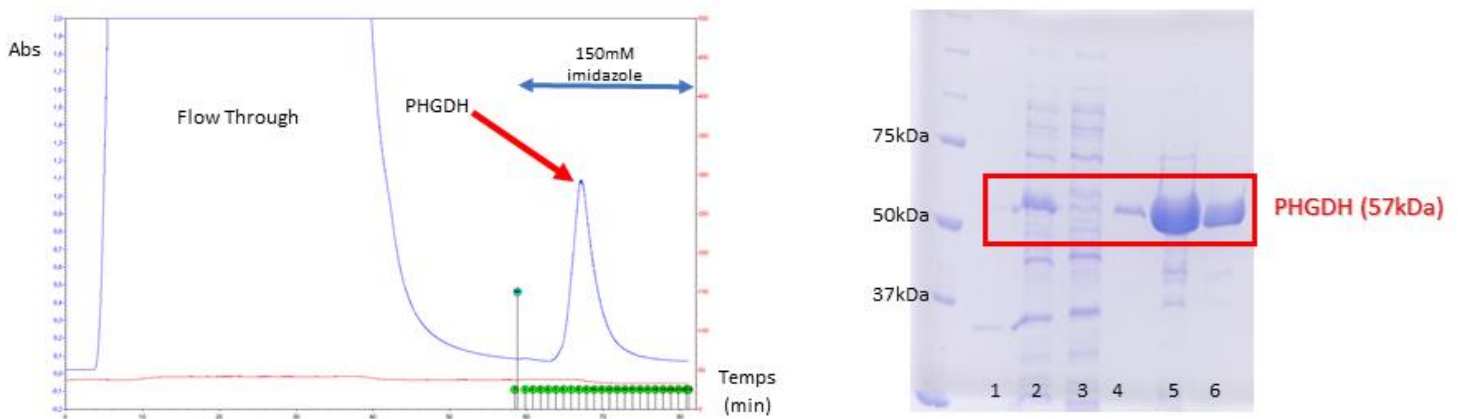


Figure 3-2: PHGDH a été purifiée par IMAC chargée au nickel (A) Chromatogramme de la purification (en rouge la conductivité en mS/cm ; en bleu l'absorbance à 280nm). (B) Gel SDS-Page de la première étape de purification : 1 culot de lyse ; 2 surnageant de lyse ; 3 Flow Through ; 4-6 pics d'élution.

Nous voyons sur la Figure 3-2 qu'après ajout de l'imidazole, qui entre en compétition avec les protéines pour la liaison du nickel, une augmentation d'absorbance correspondant en majeure partie à l'enzyme PHGDH est observée. La pureté de l'échantillon obtenu n'étant pas suffisante l'étape de cristallisation de l'enzyme, une seconde chromatographie va être appliquée. Une chromatographie d'exclusion de taille a permis d'isoler la forme active (tétramérique) de l'enzyme.

Après dialyse, afin d'éliminer l'imidazole contenu dans l'échantillon, celui-ci est chargé et élué par une chromatographie d'exclusion de taille (Superdex 200 ®) (Figure 3-3).

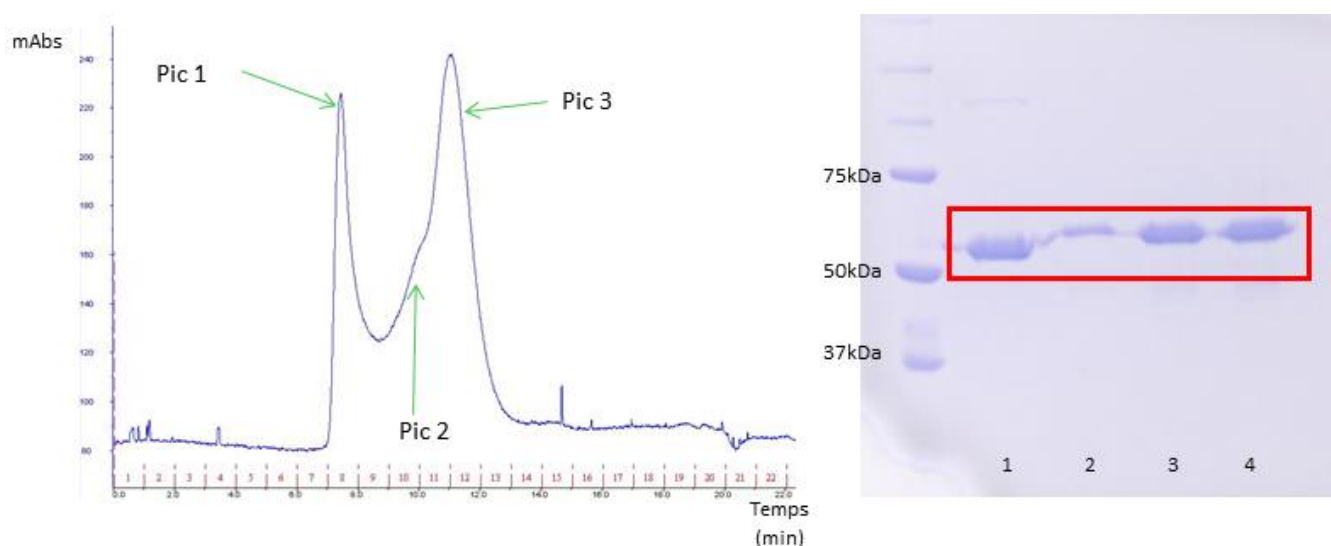


Figure 3-3: PHGDH a été purifiée par Superdex 200®. (A) Chromatogramme de la purification (en bleu l'absorbance à 280nm). (B) Gel SDS-Page de la seconde étape de purification : 1 échantillon après dialyse ; 2 pic 1 ; 3 pic 2 ; 4 pic 3.

Nous observons 3 pics différents correspondant à l'enzyme PHGDH. A l'aide d'une droite de calibration, nous avons déterminé le nombre de monomères correspondant à chaque pic. Le pic 1, de plus haut poids moléculaire, correspond au volume mort de la colonne. Nous supposons qu'il s'agit d'un agrégat de l'enzyme. Le pic 2 et 3 représentent respectivement la forme tétramérique et dimérique de PHGDH.

Afin de s'assurer que l'enzyme produite est active, nous avons testé son activité à l'aide d'un test de fluorescence. Nous avons mesuré l'apparition du NADH qui résulte de la

réduction du NAD⁺ afin de permettre la transformation du 3-PG en PPyr au cours du temps (Figure 3-4).

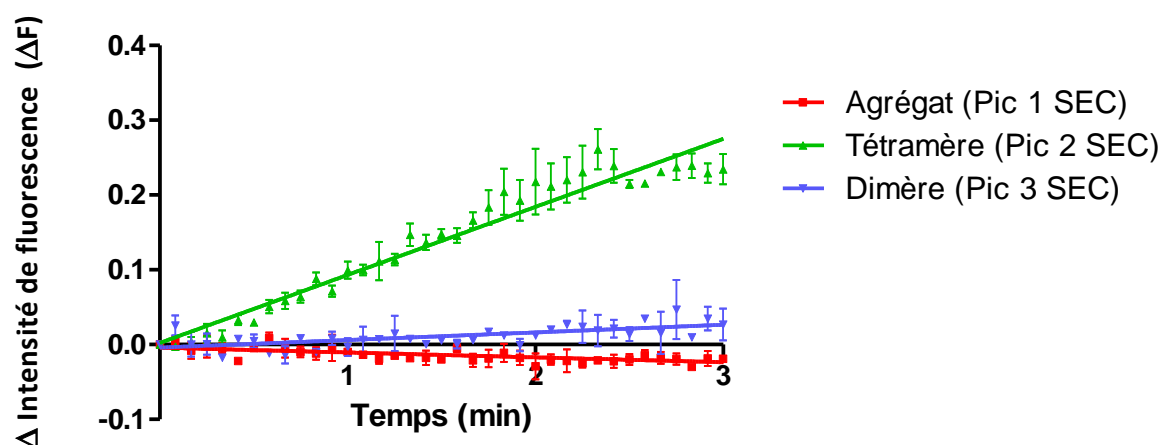


Figure 3-4: Evolution de l'intensité de fluorescence correspondant à l'apparition du NADH au cours du temps pour les 3 échantillons de PHGDH obtenus après la chromatographie d'exclusion de taille.

Nous observons que seule la forme tétramérique de l'enzyme présente une activité. Au terme de cette purification nous avons réussi à obtenir jusqu'à 38mg de protéine par litre de culture. Le lot d'enzyme pour la cristallisation a été concentré à 3,5 ; 5.2 et 11.6mg/mL. Au-delà de cette concentration la protéine commençait à précipiter. Les échantillons concentrés ont ensuite été conservés dans un tampon 50mM Tris (8,5), 10mM MgCl₂, 250mM NaCl, 20% glycérol à -80°C. Afin de trouver des conditions favorables à la cristallisation de l'enzyme nous avons criblé les kits Hampton Research 1 et 2 (HR2-110 ; HR2-112) et Wizard CRYO 1 et 2 (1009536 ; 1009537) à une concentration de 3,5 et 11,6mg/ml avec un ratio goutte/réservoir 1/1 et 1/2 dans une plaque 96 puits. L'échantillon a été incubé avec 5mM en NAD⁺ avant chaque essai de cristallisation.

Les conditions prometteuses ont été reproduites dans une plaque 24 puits afin de trouver les conditions de cristallisation optimales. La Figure 3-5 présente une partie des cristaux obtenus. Au total, plus de 1200 conditions faisant varier la concentration en enzyme, le

ratio protéine/réservoir, le pH, la concentration en sel et en agent précipitant ont été testées.

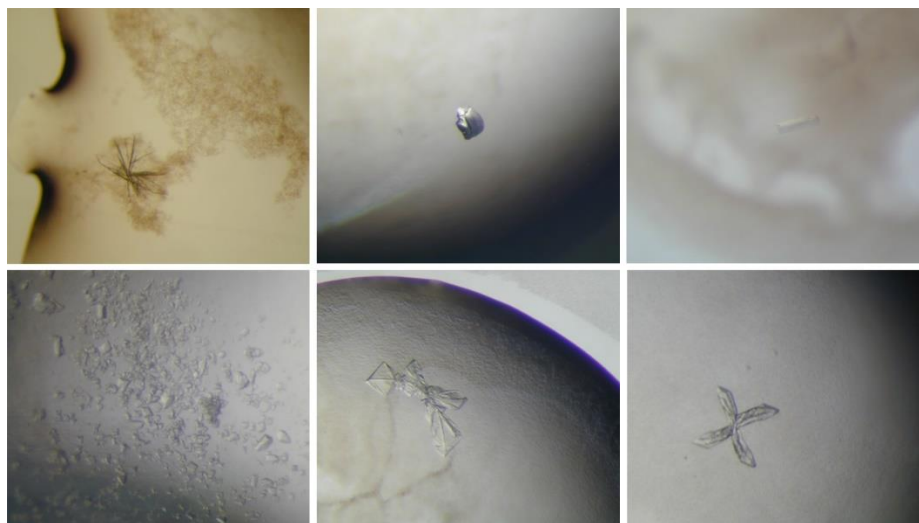
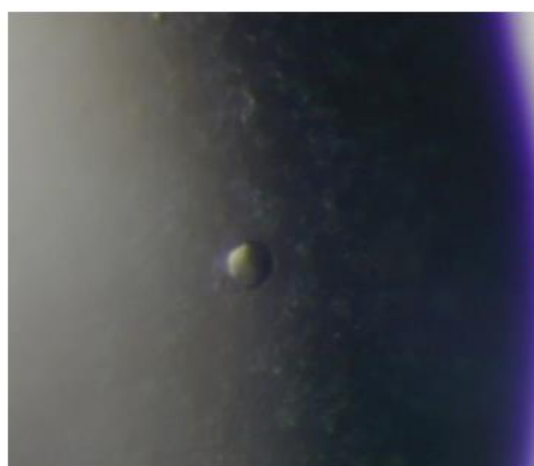
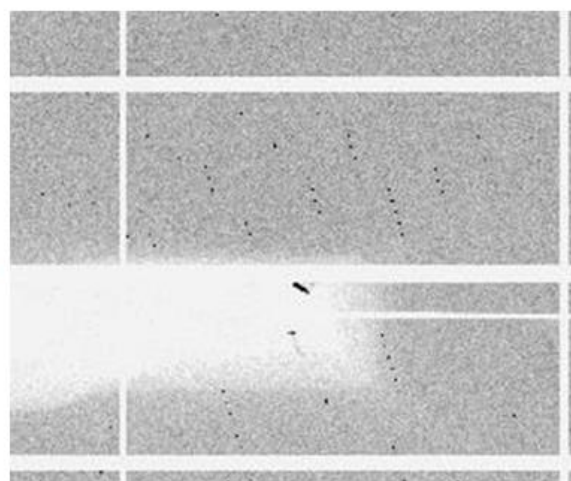


Figure 3-5: Cristaux obtenus lors de la cristallo-genèse de PHGDH sauvage.

Les cristaux obtenus ont été testés au synchrotron SOLEIL (Gif-sur-Yvette) afin de vérifier s'il s'agit bien de cristaux de protéine et si la diffraction est assez bonne pour résoudre la structure de PHGDH. Malheureusement, la plupart des cristaux obtenus étaient des cristaux de sels ou ne diffractaient pas. Un seul cristal (Figure 3-6) présentait un pattern de diffraction correspondant à une protéine. La résolution du jeu de données étant de 8Å nous n'avons pas pu générer une carte de densité électronique. Les coefficients de Matthews qui permettent sur base du volume de l'unité asymétrique de prédire le nombre de monomères dans celle-ci suggère que huit monomères sont présents (Figure 3-6). Il est par conséquent possible d'avoir deux formes tétramériques au sein de l'unité asymétrique.



0.1mM $\text{NaH}_2\text{C}_6\text{H}_5\text{O}_7$ (citrate de sodium) pH=5.6
 0.2M $\text{KNaC}_4\text{H}_4\text{O}_6$ (Potassium sodium de tartrate)
 1.8M $(\text{NH}_4)_2\text{SO}_4$
 3.5mg/mL of His- ϕ PHGDH



Unité de la cellule
 a, b, c (Å) 312,80, 312,80, 157,86
 α, β, γ (°) 90,00; 90,00; 120,0
 Resolution 8Å 49,5-7,99
 Groupe d'espace P 6 2 2 (No. 177)

Nmol/ Unité	Coef. Matthews	% solvant	P (7.99)	P(tot)
6	3.24	62.04	0.10	0.07
7	2.78	55.71	0.22	0.19
8	2.43	49.38	0.32	0.32
9	2.16	43.06	0.24	0.29
10	1.94	36.73	0.07	0.10

Figure 3-6: Cristallo-genèse, diffraction des rayons X et résultats extraits de PHGDH sous forme sauvage. (A) Photo du cristal et conditions de cristallisation ; (B) Pattern de diffraction et paramètre de l'unité asymétrique ; (C) Tableau reprenant les coefficients de Matthews et les valeurs associées.

Malgré de nombreuses tentatives de recristallisation nous n'avons jamais pu réobtenir un cristal dans ces conditions. Etant donné la difficulté d'obtenir des cristaux de PHGDH sous forme sauvage plusieurs approches ont été envisagées : fusionner l'enzyme à une protéine afin d'augmenter sa solubilité et favoriser la cristallisation ; réduire l'entropie en surface de l'enzyme afin de faciliter les contacts inter-protéines ; travailler avec l'enzyme tronquée qui a déjà été cristallisée.

3.2 PHGDH fusionné à une Maltose Binding Protein (MBP)

L'ajout d'une protéine de fusion telle qu'une Maltose Binding Protein (MBP) est une technique permettant d'augmenter l'expression et la solubilité et d'améliorer la purification de l'enzyme à laquelle elle est fusionnée⁷¹. Cette technique permet la cristallisation d'un nombre grandissant de protéines chaque année⁷².

Le plasmide codant pour l'enzyme PHGDH fusionnée à une MBP nous a été fourni par le laboratoire du prof. Olivier Féron (UCL). Il contient la séquence codante pour PHGDH insérée dans un vecteur pVP16. Ce plasmide permet l'expression de l'enzyme avec un His-tag et une MBP en N-terminal (His₈ – MBP – PHGDH) régulée par un T7 promoteur (Figure 3-7). Le vecteur contient en plus un gène de résistance à l'ampicilline.

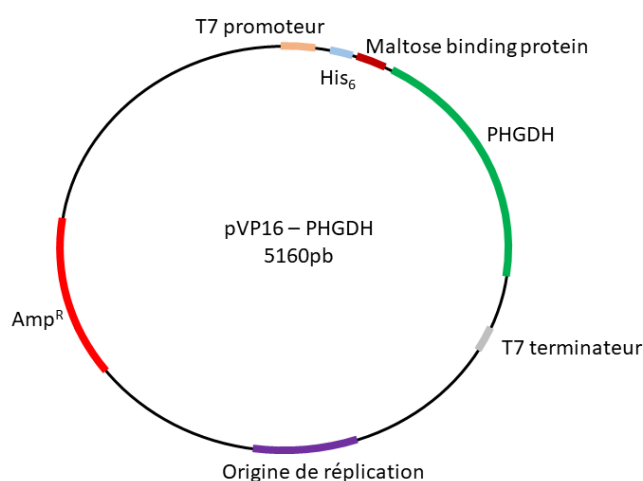


Figure 3-7: Schéma représentant le plasmide constitué du vecteur pVP16 et le gène codant pour PHGDH sous sa forme complète.

Le protocole d'expression et de purification a été mis au point et optimisé. Seuls les résultats obtenus pour le protocole optimisé seront présentés.

Le plasmide a été inséré par thermotransformation dans des bactéries *E. coli* BL21 (DE3). Les bactéries transformées ont été sélectionnées grâce à l'ampicilline et cultivées dans un milieu LB liquide. L'expression de la protéine d'intérêt a été induite par ajout de 0.5mM IPTG et incubée à 18°C durant 18heures. Les bactéries ont été lysées par sonication afin de récupérer la protéine dans la fraction soluble.

La purification de la protéine a été réalisée grâce aux *tags* présents sur celle-ci (His et MBP). L'échantillon a été chargé sur une colonne d'affinité au nickel (HisTrap FF crude, GE healthcare) branchée en série avec une colonne d'affinité à la *maltose binding protein* (MBPTrap HP, GE Healthcare). De cette façon, la protéine s'accroche sur l'IMAC grâce à son His-tag, ce qui permet d'éliminer la majorité des impuretés. L'échantillon est ensuite élué avec de l'imidazole et la protéine s'accroche sur la MBPTrap alors que les impuretés ayant une affinité pour l'IMAC sont éliminées. Finalement, PHGDH pure a été élué avec un tampon contenant 2 mM maltose. Le chromatogramme de purification ainsi que les gels SDS-Page associés sont présentés à la Figure 3-8.

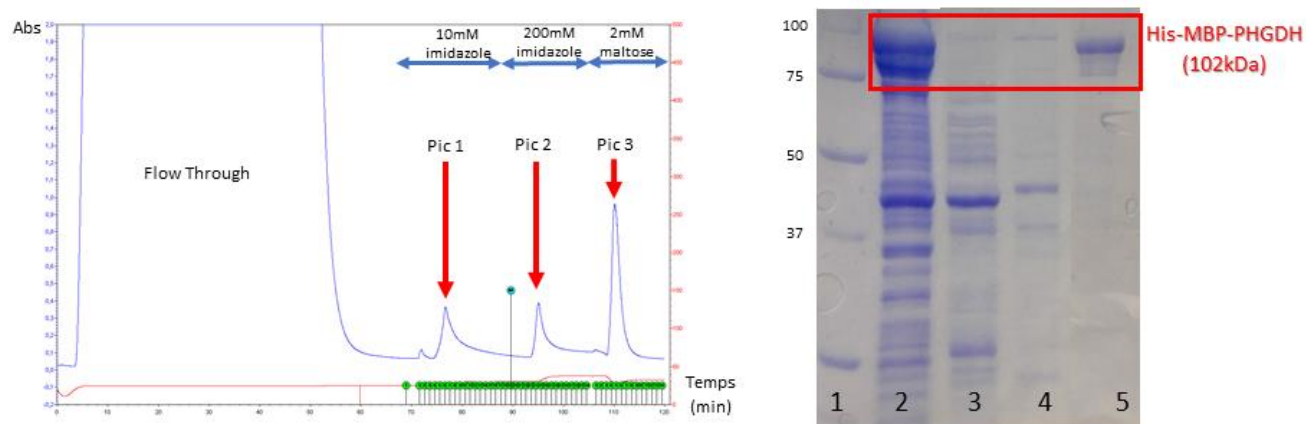


Figure 3-8: His-MBP-PHGDH a été purifiée par IMAC chargée au nickel suivie par une MBPTrap. (A) Chromatogramme de la purification (en rouge la conductivité en mS/cm; en bleu l'absorbance à 280nm). (B) Gel SDS-Page de la purification : 1 marqueur de poids moléculaire ; 2 flow through ; 3 Pic 1 ; 4 Pic 2 ; 5 Pic 3.

Nous observons sur la Figure 3-8 qu'après ajout de l'imidazole la majorité des protéines éluées sont des impuretés, alors qu'après ajout du maltose nous récupérons un échantillon de PHGDH de pureté suffisante pour la cristallisation. L'activité de chaque pic a été testée et comme nous pouvons le voir à la Figure 3-9, le pic 3 est le plus actif (remarquons une différence d'échelle en comparaison avec la Figure 3-4 qui est due à l'utilisation d'un spectromètre différent).

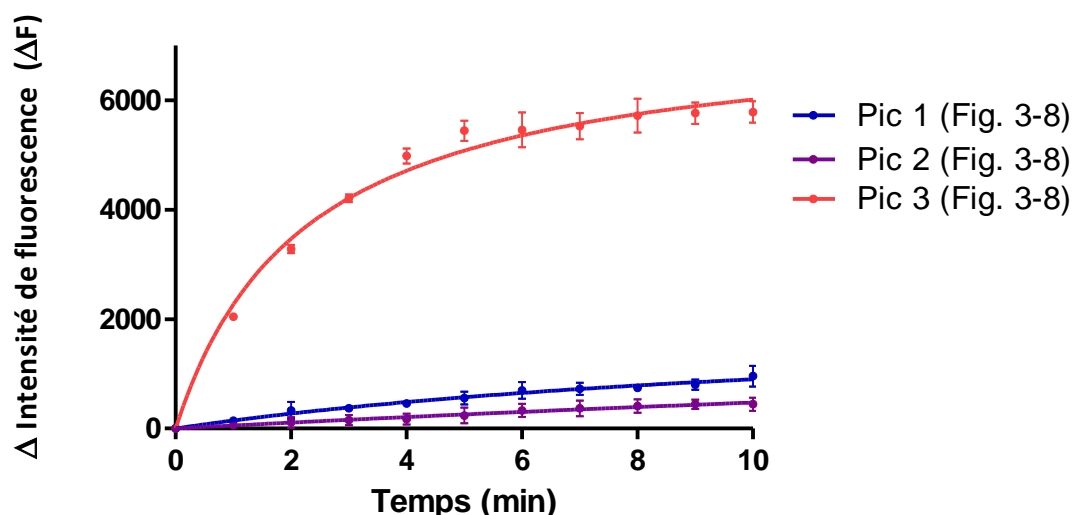


Figure 3-9: Evolution de l'intensité de fluorescence correspondant à l'apparition du NADH au cours du temps pour les 3 échantillons après purification de His-MBP-PHGDH.

L'enzyme a pu être concentrée jusqu'à 30mg/mL sans observer de signe de précipitation montrant l'impact de la MBP sur la solubilité du complexe. L'échantillon a été conservé dans un tampon composé de 50mM Tris (7.5), 100mM NaCl, 200mM imidazole, 2mM maltose et 20% glycérol à -80°C. Plusieurs concentrations en enzyme ont été testées : 5, 18 et 25 mg/mL. Après un premier criblage des conditions du *kit* Hampton Research 1 et 2 (HR2-110 ; HR2-112), les conditions prometteuses ont été reproduites et affinées en plaques 24puits. Malgré un criblage de plus de 1000 conditions, nous n'avons jamais pu obtenir de cristaux de protéine.

3.3 Réduction d'entropie de surface sur l'enzyme PHGDH complète

La réduction d'entropie de surface (Surface Entropy Reduction, SER) est utilisée depuis plusieurs années afin de faciliter la cristallisation de protéines⁷³⁻⁷⁵. Celle-ci permet de favoriser les contacts inter-protéines et par conséquent de faciliter leur cristallisation. Cette approche se base sur le fait que dans les réseaux cristallins, les valeurs d'enthalpie des forces intermoléculaires sont faibles et que de ce fait la cristallisation est sensible aux changements entropiques. Il a aussi été observé que des protéines récalcitrantes à la cristallisation comprennent des boucliers entropiques formé par des chaînes latérales polaires longues et flexibles. Un webserveur a été mis en place par l'université de Californie à Los Angeles (UCLA) afin de proposer des mutations possibles pour réduire l'entropie en surface de la protéine. Le logiciel se base sur 3 paramètres pour déterminer les clusters d'acides aminés les plus prometteurs pour la SER :

- La prédiction de la structure secondaire va permettre d'identifier les « *coil regions* ». Les résidus présents dans ces régions sont considérés comme favorables pour la SER car ces régions se trouvent généralement en surface de la protéine.
- Le profil entropique pour la séquence complète est calculé grâce à la table des entropies des chaînes latérales de Sternber⁷⁶. Les régions présentant un haut profil entropique sont de bons candidats pour la SER.
- La comparaison de la séquence de la protéine à celle d'autres protéines déjà cristallisées permet d'identifier les séquences semblables possédant des acides aminés mutés. Ceux-ci seront favorable à la SER. Dans le cas inverse, ils y seront défavorables.

Cette méthode permet de proposer des acides aminés à muter même si aucune information structurale à propos de l'enzyme n'est connue. Les 3 clusters les plus favorables et présents en surface de la protéine ont été sélectionnés (Figure 3-2Figure

3-10). Par soucis de lisibilité, le cluster #1 sera appelé mutant 1, cluster #2 mutant 2 et le cluster #3 mutant 3.

Mutations:

Cluster #1: Résidus 134 - 137: ASMKDGKW**ERKK**

- E 134 → A
- K 136 → A
- K 137 → A

Cluster #2: Résidus 32 - 34: **EKQ**

- E 32 → A
- K 33 → A
- Q 34 → A

Cluster #3: Résidus 67 - 69: AA**EKLQ**

- E 68 → A
- K 69 → A

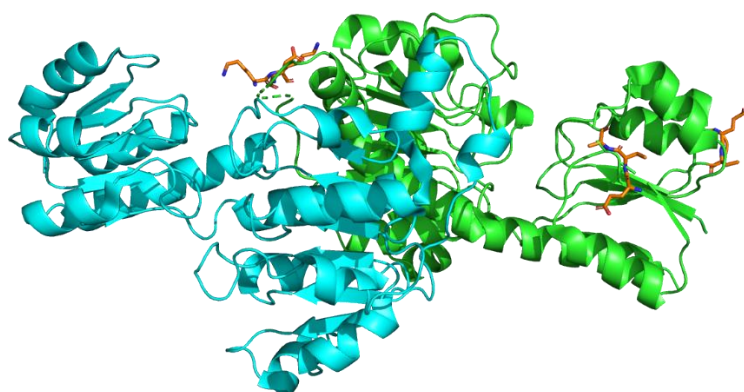


Figure 3-10: Détail et visualisation des mutations effectuées sur PHGDH. (A) Récapitulatif des mutations envisagées pour réduire l'entropie en surface de PHGDH. (B) Visualisation des 3 clusters de mutations sur la structure de PHGDH disponible (pdb : 2G76) (en cyan et vert les deux monomères contenus dans l'unité asymétrique, en orange les acides aminés mutés).

Les gènes mutés codant pour l'enzyme ont été commandés chez Integrated DNA Technologies. La séquence en base nucléotidique a été optimisée afin de conserver les paires de base les plus fréquentes chez *E. coli*. Le gène a été préalablement amplifié et inséré dans un vecteur pET-28a. Les étapes d'expression et de purification sont identiques à celles préalablement décrites pour l'enzyme sauvage. Les mutants 1 et 2 ont montré une très faible expression rendant l'utilisation de la chromatographie d'exclusion de taille impossible à cause des pertes inhérentes à la faible proportion

d'enzyme active. La modification des paires de bases nucléotidiques n'a donc pas permis une meilleure expression de la protéine. La Figure 3-11 représente le processus de purification pour le mutant 3.

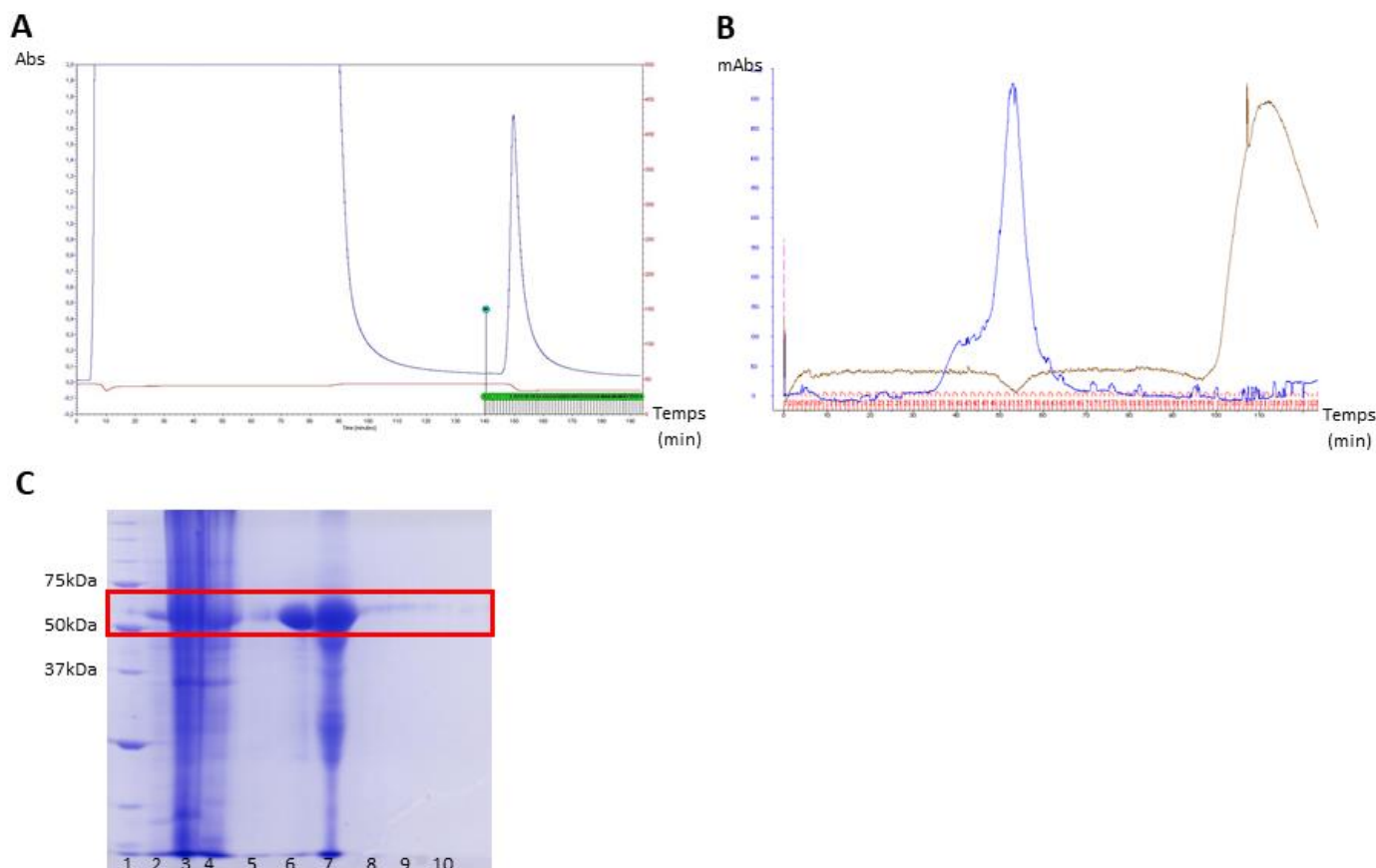


Figure 3-11: Le mutant E68A K69A a été purifié par IMAC chargée au nickel suivie par une exclusion de taille Superdex 200™. (A) Chromatogramme de la purification par IMAC (en rouge la conductivité en mS/cm ; en bleu l'absorbance à 280nm). (B) Chromatogramme de la purification par Superdex 200™ (en rouge la conductivité en mS/cm ; en bleu l'absorbance à 280nm) (C) Gel SDS-Page de la purification : 1 marqueur de poids moléculaire ; 2 milieu induit ; 3 culot de lyse ; 4 surnageant de lyse ; 5 Flow through IMAC ; 6 PHGDH après IMAC ; 7 précipité après dialyse ; 8 S200 pic 1 ; 9 S200 pic 2 ; 10 S200 pic 3.

Pour le mutant 3, nous observons une bonne expression de l'enzyme. Celle-ci a donc pu être purifiée par chromatographie d'exclusion de taille. Durant la purification nous avons observé une instabilité de l'échantillon. Pendant la dialyse réalisée entre l'IMAC et la Superdex 200™ qui permet d'éliminer l'imidazole et de réduire la concentration en sel, l'enzyme a précipité (Figure 3-11C puit 7). Cette précipitation a aussi été observée après purification par exclusion de taille. La mutation de glutamates en surface de la protéine par des alanines peut en effet réduire sa solubilité et c'est ce que nous observons ici avec une plus grande instabilité de notre échantillon. Après purification

par exclusion de taille nous avons pu récupérer jusqu'à 21mg de PHGDH mutés. L'échantillon a été conservé à -80°C dans un tampon 50mM Tris (8,5), 10mM MgCl₂, 250mM NaCl, 20% glycérol.

Au terme de ces purifications, les trois mutants ont pu être exprimés et purifiés. Les mutants 1 et 2 ont été purifiés par IMAC à cause de leur faible expression et le mutant 3 a subi une IMAC suivie d'une exclusion de taille. Fait intéressant, le mutant 1 est complètement inactif alors que les deux autres mutants ont conservé leur activité (l'inactivité du mutant 1 est discutée plus en détail dans le chapitre 3).

Pour chaque mutant, nous avons criblé les conditions du kit Hampton Research 1 et 2 (HR2-110 ; HR2-112) à des concentrations de 3,5 et 5mg/mL de protéines dans des plaques 96puits. De faibles concentrations en enzyme ont été utilisées étant donné la faible solubilité de celle-ci. Les conditions les plus prometteuses ont été reproduites et affinées mais nous n'avons jamais pu obtenir de cristaux de protéine adaptés pour le synchrotron. Au total, 350 conditions auront été testées pour ces mutants.

3.4 Expression et purification de l'enzyme tronquée (3-338)

Les structures cristallographiques présentes sur la Protein Data Bank ayant toutes été obtenues avec une forme tronquée de PHGDH, nous avons décidé de travailler avec celle-ci afin de générer des jeux de données exploitables. Le plasmide a été acheté chez Addgene (originellement offert par Nicola Burgess-Brown).

Le plasmide contient une séquence codante pour l'enzyme tronquée (3 – 338) qui a été insérée sur un vecteur pNIC28-Bsa4. Ce vecteur contient un gène de résistance à la kanamycine et insère un His₆-tag clivable à la TEV protéase à l'extrémité N-terminale (Figure 3-12). L'expression de l'enzyme peut être induite avec de l'IPTG grâce à une régulation par un promoteur T7.

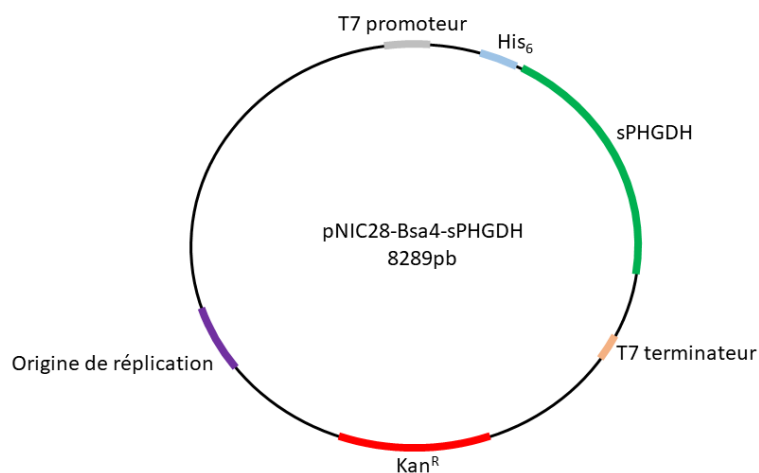


Figure 3-12: Schéma représentant le plasmide constitué du vecteur pNIC28-Bsa4 et le gène codant pour PHGDH sous sa forme tronquée (SPHGDH).

Le protocole d'expression et de purification nous a été fourni par Dr J. Unterlass. Une chromatographie d'affinité permet de lier l'His-tag de la protéine et d'éliminer une majorité des impuretés. Celle-ci est suivie par le clivage du tag et un second passage sur l'IMAC qui laisse passer la protéine clivée mais interagit avec le tag clivé, l'enzyme non clivée et les impuretés restantes. Cette seconde purification a permis d'obtenir une protéine sans *tag* et pure. Finalement l'échantillon est élué sur une chromatographie échangeuse d'anion (HiTrap Q HP™, GE Healthcare) afin de s'assurer de l'obtention d'un échantillon homogène. Un traitement au charbon actif a été appliqué sur certains échantillons afin d'éliminer le NAD⁺ présent dans la poche du cofacteur. Les étapes de purification de la protéine sont présentées à la Figure 3-13.

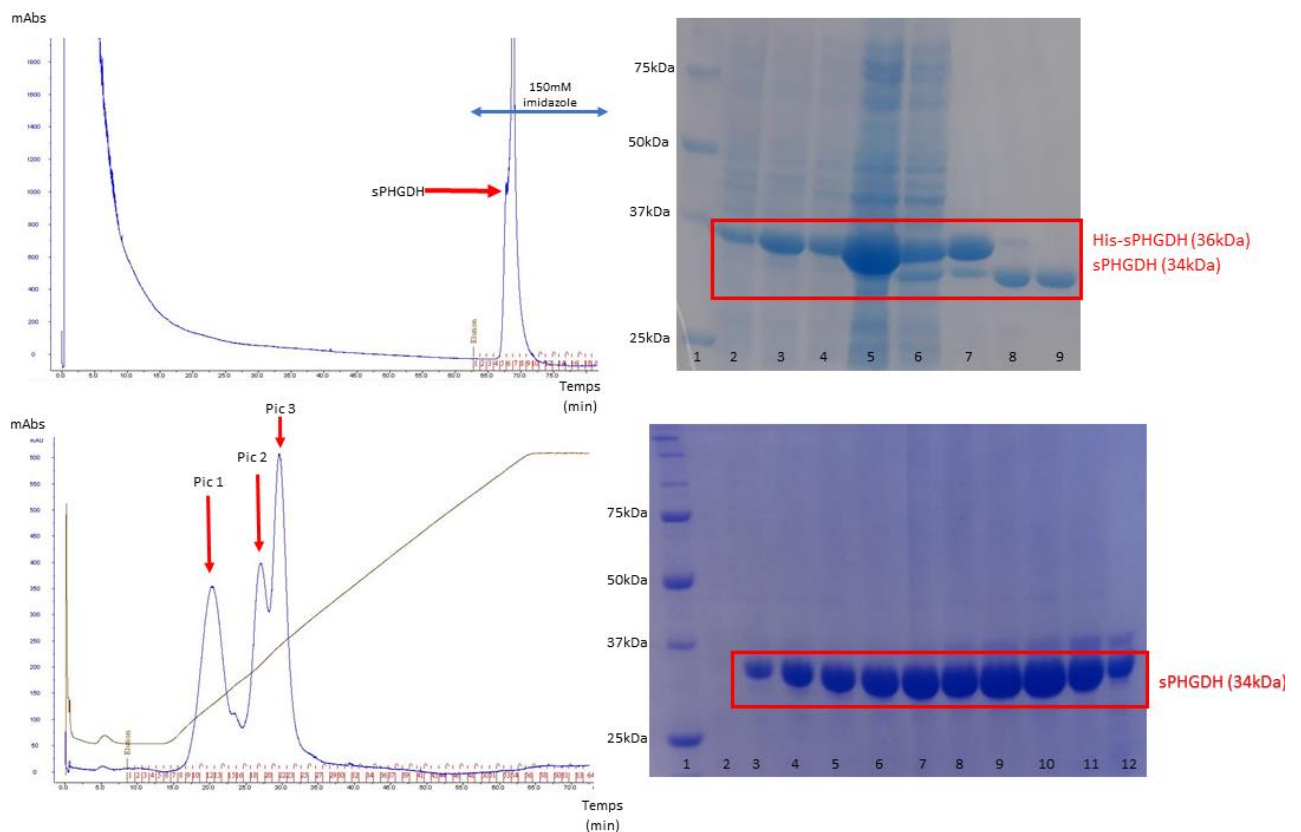


Figure 3-13: sPHGDH a été purifiée par IMAC chargée au nickel suivie par une seconde IMAC après clivage de l'His-tag et a finalement été purifiée par chromatographie échangeuse d'anion. (A) Chromatogramme de la purification par IMAC (en bleu l'absorbance à 280nm). (B) Gel SDS-Page des deux IMAC : 1 marqueur de poids moléculaire ; 2 milieu de culture ; 3 milieu de culture après induction ; 4 culot de lyse ; 5 surnageant de lyse ; 6 flow through première IMAC ; 7 élution première IMAC ; 8 échantillon après clivage ; 9 sPHGDH clivé après deuxième IMAC. (C) Chromatogramme de la purification sur colonne échangeuse d'anion (en bleu l'absorbance à 280nm ; en brun la conductivité en mS/cm). Gel SDS-Page de la chromatographie échangeuse d'anion : 1 marqueur de poids moléculaire ; 2-5 pic 1 ; 6-9 pic 2 ; 10-12 pic 3.

Nous observons qu'après les deux premières étapes de purification et le clivage nous obtenons l'enzyme correctement clivée et pure. La chromatographie par échangeuse d'ions indique que nous avons un échantillon hétérogène de sPHGDH. En effet, trois pics distincts sont élués avec le gradient en sel (0 à 600mM NaCl en 50minutes). La première hypothèse était qu'il s'agissait de différents états d'oligomérisation de la protéine. Toutefois les trois pics éluent en même temps en chromatographie d'exclusion de taille, attestant que le volume global de ces trois populations est identique. L'hypothèse a donc été faite qu'il s'agit de changements en surface de la protéine modifiant les interactions que celle-ci peut faire avec la colonne. Nous avons tenté de cristalliser les trois populations et avons pu obtenir des cristaux avec les populations correspondant aux pics 1 et 2.

Une partie de la protéine produite a subi un traitement au charbon actif afin d'éliminer le NAD⁺ présent dans le site actif alors que le reste a conservé son cofacteur. Les deux échantillons ont été conservés dans un tampon 25mM HEPES (pH=7.5), 100mM NaCl, 0.5mM TCEP à -80°C.

Les premiers essais de cristallisation réalisés à l'institut Karolinska (Suède) sous la supervision du Dr. Judith Unterlass ont criblé une série de kit commerciaux (PACT, Index et JCSG+) en présence d'une série d'inhibiteurs fournis par nos collaborateurs (UCL- laboratoire du Prof. Raphaël Frédérick). Ce premier criblage a montré que le bis-tris (pH=5.5 - 7.5) et le PEG 3350 sont favorables à la cristallisation de sPHGDH.

Sur base de cette information, un criblage rationnel des conditions de cristallisation a été réalisé afin de trouver les meilleures conditions pour la production de cristaux permettant une bonne diffraction des rayons X. Il a été montré que bien que le bis-tris soit le meilleur tampon de cristallisation, des cristaux peuvent être obtenus tant que le pH est compris entre 5.5 et 7.5. De cette façon des cristaux ont été produits en présence de MES et HEPES alors que nous n'en avons jamais observé à des pHs plus élevés (bicine et CHES). Le PEG 3350 est le seul agent précipitant avec lequel nous ayons obtenus des cristaux. Le sel de base utilisé est le Li₂SO₄. Nous avons testé différentes combinaisons de cation/anion qui ont montré une faible importance du cation. Nous obtenons des cristaux avec différents cations mono et bivalents contrairement à l'anion où le sulfate semble être essentiel. En effet, l'utilisation d'un chlorure ne permet pas la formation de cristaux alors que l'acétate ne génère des cristaux que dans certaines conditions. sPHGDH a aussi cristallisé en absence de sel.

Ce criblage a permis de dire que pour cristalliser sPHGDH il faut : du PEG 3350, un sel composé de sulfate (ou pas de sel) et un pH compris entre 5,5 et 7,0. Le tampon et le cation du sel ne semblent pas affecter significativement la capacité de l'enzyme à cristalliser et il est possible de changer l'anion du sel à condition de conserver un groupement relativement grand. Toutefois, pour ce dernier critère, un échantillon plus vaste de sel serait nécessaire afin de pouvoir tirer des conclusions. La Figure 3-14 montre un échantillon des différentes formes de cristaux obtenus. Au total, plus de 4300 conditions ont été testées.

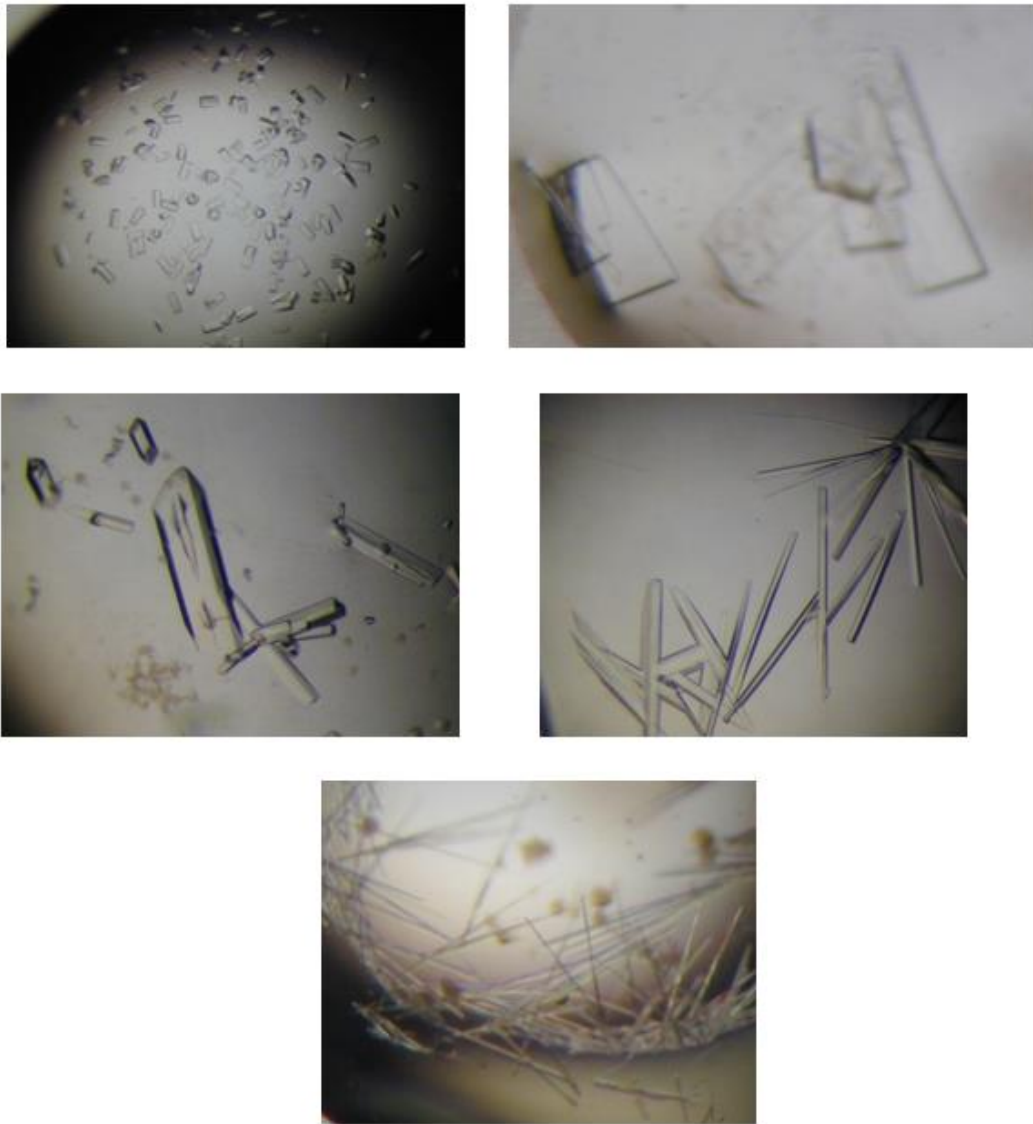


Figure 3-14: Cristaux obtenus lors de la cristallogénèse de PHGDH tronquée.

Les cristaux obtenus ont été analysés au synchrotron de SOLEIL (ligne PROXIMA 1 ET 2) confirmant qu'il s'agissait bien de cristaux de protéine. La plupart de ceux-ci ont cristallisé dans un groupe d'espace $P 2_1 2_1 2$ (18). Le Tableau 3-1 reprend les données cristallographiques des groupes d'espace $P 2_1 2_1 2$ pour la meilleure résolution obtenue de chaque ligand co-cristallisé. Toutes les données obtenues dans le groupe d'espace $P 2_1 2_1 2$ présentent 4 monomères dans l'unité asymétrique selon les coefficients de Matthews.

Tableau 3-1: Liste des données cristallographiques et résolutions en fonction du ligand co-cristallisé avec sPHGDH.

LIGAND	CELLULE						GROUPE D'ESPACE	RESOLUTION (Å)
	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)		
Disulfirame	91.68	287.9	52.62	90	90	90	P 2 ₁ 2 ₁ 2	3.03
CMFA-005	91.31	285.28	52	90	90	90	P 2 ₁ 2 ₁ 2	2.71
Apo-enzyme	91.91	286.81	52.42	90	90	90	P 2 ₁ 2 ₁ 2	3.26
NAD ⁺ + 3PG	91.23	283.76	51.93	90	90	90	P 2 ₁ 2 ₁ 2	2.99

Dans l'unité asymétrique, nous observons deux dimères indépendants de sPHGDH. Bien que quatre monomères soient présents dans l'unité asymétrique ceux-ci n'ont jamais permis d'obtenir la forme tétramérique de l'enzyme confirmant les résultats présentés dans la littérature et confirmant le fait que le domaine catalytique n'est pas suffisant pour obtenir PHGDH sous forme tétramérique. Malgré les efforts investis dans l'affinement de ces jeux de données, nous n'avons jamais pu obtenir un Rwork/Rfree inférieur à 30%. Dans chacun des dimères nous observons un *substrate binding domain* pour lequel nous n'arrivons pas à générer de carte densité (Figure 3-15). En effet, le *SBD* est décrit comme étant flexible⁵⁵ et l'hypothèse est faite que suite à cette flexibilité nous ne pouvons obtenir une bonne carte de densité pour cette partie de la protéine.

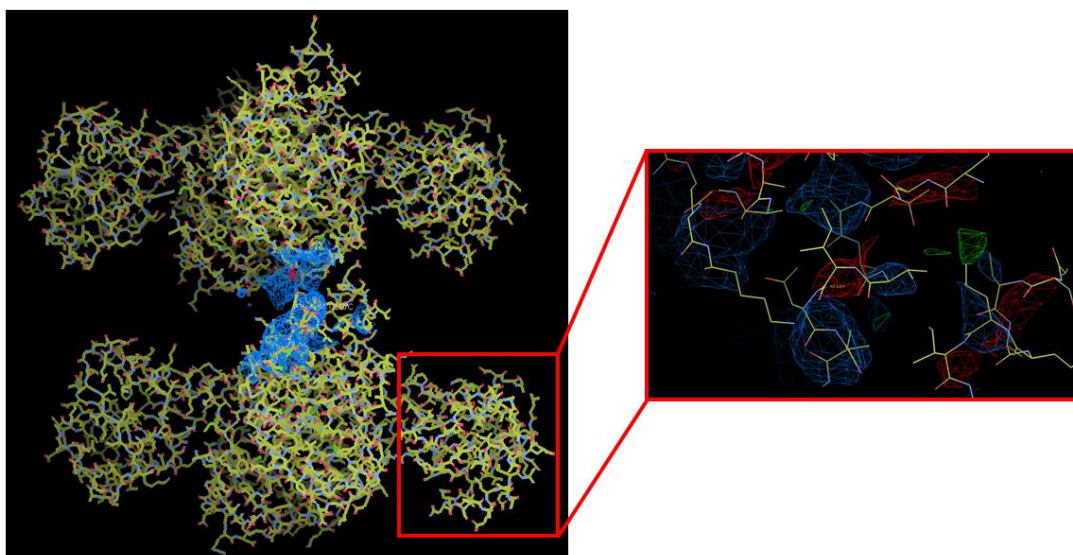


Figure 3-15: Carte de densité et coordonnées générées pour sPHGDH cristallisé dans un groupe d'espace $P 2_1 2_1 2$, un zoom a été fait sur le SBD.

Le deuxième groupe d'espace observé est un groupe $P 1$ (le détail de ces jeux de données est décrit aux section 5 et 6 lors de la présentation de ces structures). Celui-ci est constitué d'un seul dimère et est donc plus proche de ce qu'il existe dans la littérature. Ces données ont été obtenues en co-cristallisant l'acide fluoropyruvique et la l-cystéine, des ligands identifiés au sein du laboratoire (détail à la section 6). Elles ont pu être affinées et les structures obtenues sont présentées au chapitre 4 en même temps que la découverte de ces composés.

3.5 Discussion

Au terme de ce travail nous avons mis au point un protocole d'expression et de purification de l'enzyme sous sa forme complète, active et pure. Après de nombreuses tentatives nous n'avons pu obtenir qu'une seule fois un cristal de protéine mais ne diffractant pas assez pour permettre de résoudre sa structure. Avec les moyens à notre disposition il semble difficile de cristalliser l'enzyme sous cette forme. C'est pourquoi nous avons mis en place plusieurs stratégies afin de favoriser la cristallisation de l'enzyme.

Le premier moyen mis en œuvre, l'ajout d'une *maltose binding protein*, a permis comme décrit dans la littérature de grandement augmenter la solubilité de la protéine tout en conservant son activité. Malheureusement, ceci n'a pas permis de favoriser la formation des cristaux d'enzyme. Nous avons ensuite muté des acides aminés en surface de la protéine afin de réduire son entropie et favoriser les contacts inter-protéines. De façon générale les mutants ont réduit la stabilité voire même l'activité pour la mutation 2 de l'enzyme sans que cela ne permette d'obtenir des cristaux de protéines. Toutefois, il aurait été intéressant de chercher des tampons stabilisant ces mutations. Faire des essais de cristallogenèse avec des échantillons ayant montré de l'instabilité dans les tampons utilisés réduit les chances d'une cristallisation réussie. La stabilité de l'enzyme dans son tampon pourrait être évaluée en mesurant sa température de fusion par *Differential Scanning Fluorimetry* (DSF)^{77,78}. Nous aurions aussi pu envisager de muter les acides aminés en arginine plutôt qu'en alanine. En effet, il a été montré que cette mutation pouvait aussi favoriser les contacts inter-protéines en réduisant la perte de solubilité⁷⁹. La réduction d'entropie de surface pourrait être investiguée plus en profondeur. Finalement, l'utilisation de l'enzyme sous forme tronquée a permis l'obtention de cristaux de protéine diffractant suffisamment pour résoudre la structure tridimensionnelle.

Nous y observons une importante flexibilité du domaine de liaison au substrat comme déjà décrit dans la littérature⁵⁵. L'ajout de ligands se liant dans le site actif a permis d'obtenir des données exploitables. En effet, on peut conclure que l'ajout de ces ligands permet de réduire la flexibilité du domaine et a permis d'obtenir les structures cristallographiques de co-cristaux. Nous observons toutefois que même après liaison d'un ligand la carte de densité du SBD reste partielle, ce qui pourrait être expliqué par le pourcentage d'occupation du ligand. Nous avons aussi pu confirmer que, malgré la présence de quatre monomères de protéine dans l'unité asymétrique, celle-ci ne forme pas de tétramère. Ces résultats confirment le fait que le domaine ASB et ACT sont nécessaires pour la tétramérisation de l'enzyme. Un de ces deux domaines au moins est délétère pour la bonne cristallisation de l'enzyme. Ils sont liés au domaine catalytique par une boucle flexible de 20 a.a. Il pourrait être envisagé de réduire cette boucle afin de rigidifier le système et favoriser sa cristallisation. Nous avons principalement muté

le domaine catalytique pour réduire l'entropie en surface de celui-ci. Avec le recul que nous avons maintenant, il semble plus judicieux de réduire l'entropie en surface des domaines ASB/ACT. En effet, favoriser les contacts protéine-protéine dans cette région que nous savons difficile à cristalliser pourrait permettre la cristallisation du système dans sa globalité. Finalement, l'expression des domaines ASB/ACT seuls est une autre façon d'obtenir des informations structurales qui permettraient de reconstruire la structure de l'enzyme complète. Cette technique éviterait les problèmes liés à la flexibilité des domaines les uns par rapport aux autres.

*4. Modèle tétramérique de l'enzyme
PHGDH complète*

4 Modèle tétramérique de l'enzyme PHGDH complète

La cristallisation de l'enzyme sous sa forme complète n'a jamais permis d'obtenir des cristaux de protéines diffractant suffisamment. Afin d'obtenir des informations structurales sur PHGDH sous sa forme tétramérique, nous avons décidé d'utiliser des méthodes calculatoires.

4.1 Génération du modèle

Pour générer un modèle il est nécessaire d'avoir un template. Nous savons que PHGDH humaine est classée dans les enzymes de type 1. Parmi les 25 structures de PHGDH présentes sur la PDB, 12 proviennent d'enzymes de type 1 et seulement trois ne sont pas une forme tronquée de la forme humaine. Ces trois structures proviennent de l'organisme *Mycobacterium tuberculosis*. Elles sont toutes les trois sous forme tétramérique et sont différenciées par le ligand co-cristallisé dans le site actif.

Il a été montré que PHGDH provenant de *Mycobacterium tuberculosis* (tbPHGDH) est active sous forme tétramérique tout comme hPHGDH⁸⁰. De plus, une identité de séquence de 32% est observée entre les enzymes des deux souches³. La Figure 4-1 montre l'alignement des séquences de la forme humaine et deux formes bactérienne⁴ provenant des organismes *Mycobacterium tuberculosis* (la forme de type 1 ayant déjà été cristallisée) et *E. coli*. (la forme de type 1 ayant déjà été cristallisée). La ligne « consensus_aa » représente les acides aminés identiques en majuscule et en gras. La ligne « consensus_ss » met en évidence les régions avec une structure secondaire conservée (h pour hélice α et e pour feuillet β). Nous pouvons voir que les acides aminés conservés ou présentant la même fonction sont répartis sur toute la séquence et que la majeure partie des structures secondaires est identique. Nous observons aussi que

³ Déterminé par alignement de séquence sur le site <https://www.uniprot.org>. Programme utilisé : CLUSTALO

⁴ Alignement réalisé à l'aide du programme PROMALS3D :
<http://prodata.swmed.edu/promals3d/promals3d.php>

l'absence du domaine ASB sur la forme provenant de *E. coli* entrainerait un manque d'informations structurales sur 150 acides aminés si cet organisme était utilisé comme modèle. Ces informations confirment que tbPHGDH et hPHGDH sont structuralement semblables.

Substrate Binding Domain (SBD)

Conservation: 9 9 9 9 99 9 9
 E._coli 1 MAKVSLEKDKIKFLLVEGVHQALESIRAG-YTNIEFHKGALDDEQLKESIRDAHFIGLRSRTHLTEDV 69
 Myco._Tub. 1 MS-----LPVVLIADKLAPSTVAALGDQV-EVRNVVD--GPDRLKLLAAVPEADALLVRSATTVDREV 59
 Human 1 MAF----ANLRKVLISDSLDPCCRKILQDGGGLQVVEKQ--NLSKEELIAELQDCEGLIVRSATKVTADV 63
 Consensus_aa: Mt.....b..ALis-.l...h...L.s.s.bh.b.....s.sc-pLb...l.-tc.l.lRS.Tpls.-V
 Consensus_ss: eeeee hhhhhhhhh eeee hhhhhhh eeee hhh

Conservation: 99 99 9 99 99 9 9 9 9 9 99 9 99 9 99 9
 E._coli 70 INAAEKLVAIGAFAGITNQVDLDAAAKRGIPVFNAAPFSNTRSAVELVIGELLLLLRGVPEANAKAHRGVW 139
 Myco._Tub. 60 LAAAPKLKIVARAGVGLDNVDVDAATARGVLVNNAPTSTNIHSAEHALALLAASRQIPADASLREHTW 129
 Human 64 INAAEKLQVVGGRAGTGVDNVDLEAATRKGILVMNTPNGNSLSAAELTCGMIMCLARQIPQATASMKDGKW 133
 Consensus_aa: lsAA.KL.hlt.hthGhspVDl-AAh.+Gl.VhNp.tN..ShAEhhtbihhh.R.lP.AsAph+c..W
 Consensus_ss: hhh eeeee hhhhhh eeeee hhh

Conservation: 9 99 9 9 9 99 9 99 9 99 9 9
 E._coli 140 NKLAAGSFEARGKKLGIIGYGHIGTQLGILAESLGMVVFYDIENKLP--LGNATQVQHLSDLLNMSDVV 207
 Myco._Tub. 130 KRSSFSGTEIFGKTVGVVGLGRIGQLVAQRIAFGAYVVAYDPYVSPARAAQLGIELLSLDDLLARADFI 199
 Human 134 ERKKFMGTENLNGKTLGILGLGRIGREVATRMQSGMKTIGYDPIISPEVSASFGVQQLPLEEIWPLCDFI 203
 Consensus_aa: p+..h.thEh.GKplGllGhG+IGpblt..h.thGhbbh.YD.b.p....h..thp.b.Lp-lhs.tDhl
 Consensus_ss: eeeee hhhhhhhhhhh eeeee hhh eee hhhhhhh ee

Nucleotide Binding Domain (NBD)

Conservation: 9 9 9 9 9 9 9 99 99 99 9 9 999 99
 E._coli 208 SLHVPENPSTKNMGAKEISLMKPGSLLINASRGTVVDIPALADALASKHLAGAAIDVFPTEPATNSDPF 277
 Myco._Tub. 200 SVHLPKTPETAGLIDKEALAKTKPGVIVNAARGGLVDEAALADAITGGHVRAAGLDVFATEPCT----- 264
 Human 204 TVHTPLLPTTGLLNDNTFAQCKKGVRVNVNARGGIVDEGALLRALQSGQCAGALDVFTEEPPEPR----- 268
 Consensus_aa: olHhPb.PpT.shhs.p.hthhK.Gs.lNntRGslVDbsALhcAl.t.ph.tAtlDVFspEPsp.....
 Consensus_ss: eee hhhh hhhhhh eeeee hhhhhhhhh eeeee

Substrate Binding Domain (SBD)

Conservation: 9 9 99 9 99 999 9 9 99
 E._coli 278 TSPLAEFDNVLTPHIGGSTQEAQENIGLEVAGKLIKYSNDGSTLSAVNFPEVSLPLHVGRLM----- 341
 Myco._Tub. 265 DSPLFELAQVVVTPHLGASTAEAQDRAGTDVAESVRLALAGEFVPDAVNVG--GGVVNEEVA PWLDLVRK 332
 Human 269 DRALVDHENVISCPHLGASTKEAQSRGEEIAVOFVDMVKGKSLTGVVNAQALTSAPSPHFKPWIGLAEA 338
 Consensus_aa: spsLh-h.pVl.hPhlGtST.EAQpphG.-lA.ph..h..s..h.shVN...s.shp....h.....
 Consensus_ss: hhhh eee hhhhhhhhhhhhhhhhhhh hhh hhhhhhhh

ASB Domain

Conservation: -----
 E._coli 333 LGVLAVLSDELFPVSLSVQVRGELAAEEVEVLRLSALRGLFSAVIE-DAVTFVNAPALAAERGVTAEICK 401
 Myco._Tub. 339 LGTLMRANAGSPKGTIQVITQGTSLKNAGNCLSPAVIVGLLKEASKQADVNLVNAKLLVKEAGLNVITSH 408
 Human
 Consensus_aa:
 Consensus_ss: hhhhhhhhhh eeeeeee h hhhhhhh eeeeeee

Conservation: ----- 9
 E._coli 342 -----HIHENRPG 349
 Myco._Tub. 402 ASESFNHRSVVDVRAVGAD-GSVVTVSGTLYGFPOLSKIVQINGRHFD--LRAQ--GINLIHYVDRPG 465
 Human 409 SPAAPGEGCGEGCLLAVALAGAPYQAVGLVQGT--TPVLOGINGAVFRPEVPLRRDLPLLLFRQTQSDPA 476
 Consensus_aa:h..scPt
 Consensus_ss: eeeeeee eeeeeee eeeeeee ee eeeee

ACT Domain

Conservation: 9 99 9
 E._coli 350 VLTALNKIPAEQGVNIAAOYLQTSACMGVYVIDIEADEDAEKALQAMKAIPGTIRARLLY-- 410
 Myco._Tub. 466 ALGKIGTLLGTAGVNIQAQQLSEDAEGPGATILLRLDQDVPDDVRTAIAAAVDAYKLEVVDLS 528
 Human 477 MLPTMIGLLAEAGVRLLSYQTSLSVS-DGETWHVMGISSLLPS--LEANKCH--VTEAFQFHF- 533
 Consensus_aa: hLs.h..lhtp.GVpl.t.bhp.st..s.hhh.h.hsp.lsp...pAh..h..hhch..h...
 Consensus_ss: hhhhhhhhhh eeeee eeeee hhhhhhhh eeeee

Figure 4-1: Alignement des séquences de hPHGDH et tbPHGDH effectué avec le programme PROMALS3D. Les acides aminés en rouge correspondent à des hélices α et en bleu à des feuillets β , la dernière ligne « consensus_ss » représente la conservation des structures secondaires dans les deux enzymes (h pour les hélices α et e pour les feuillets β). La ligne « consensus_aa » présente un acide aminé écrit en gras et majuscule quand celui-ci est conservé et une lettre minuscule correspondant à la fonction conservée (aliphatique: l ; aromatique : @ ; hydrophobique: h ; alcool: o ; résidu polaire: p ; minuscule: t; petit : s ; résidu encombrant: b ; chargé positivement : + ; chargé négativement: -; chargé: c.). La première ligne « conservation » reprend avec la valeur de l'index de conservation (utilisé dans les cas de l'alignement de 3 séquences ou plus), une valeur de 9 correspond à une conservation de l'acides aminés dans chaque séquence.

Une partie de la forme humaine de l'enzyme ayant déjà été cristallisée, nous avons aligné la structure de cette enzyme (sPHGDH, 2G76, conformation fermée) avec celle de *Mycobacterium tuberculosis* afin de confirmer la ressemblance structurale (Figure 4-2).

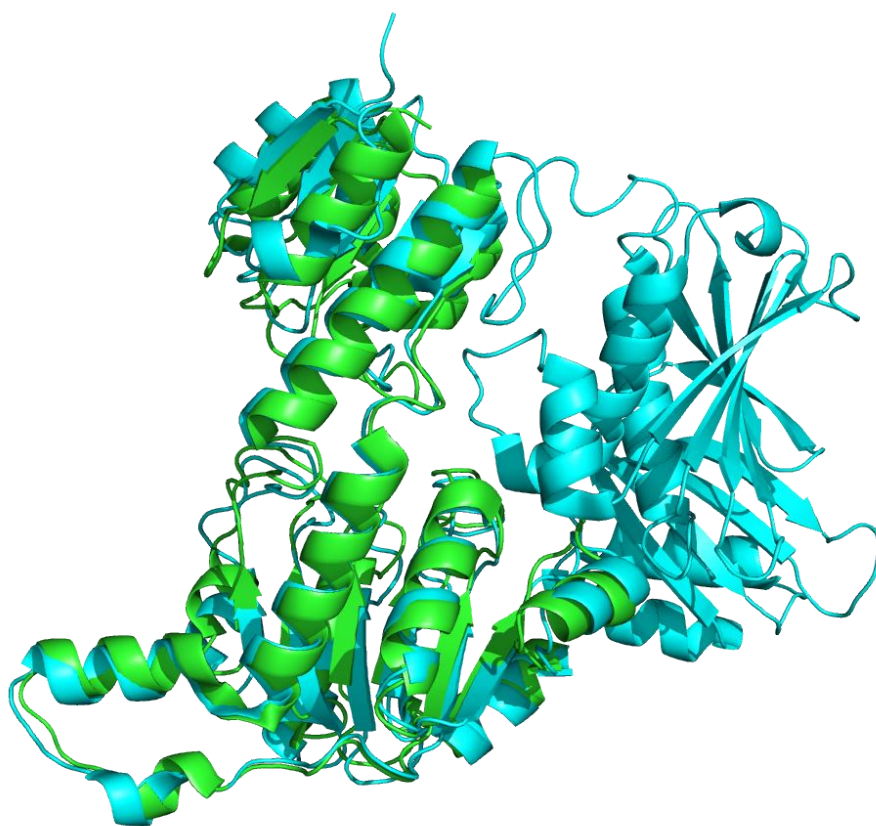


Figure 4-2: Alignement structurale de sPHGDH (vert, 2G76) et tbPHGDH (cyan, 3DDN).

Nous pouvons voir que la forme tronquée et fermée de l'enzyme humaine (vert) et la forme bactérienne (cyan) possèdent la même structure. Un *Root Mean Square Deviation*

(RMSD) de 2,2 Å est calculé entre les carbones alphas des acides aminés communs aux deux structures. Les différences entre les structures générées et la structure cristallographique sont majoritairement observées au niveau des boucles flexibles. Seule l'hélice alpha composée des acides aminés D12 jusqu'à L27 est modélisée sous forme de boucle dans nos modèles. Nous y observons un RMSD de 2.5 Å entre les modèles générés et la structure cristallographique.

Ces informations indiquent que PHGDH provenant de la souche de *Mycobacterium tuberculosis* est un bon *template* pour nos calculs.

Les webserveurs SWISS MODEL et i-TASSER ont été utilisés pour la modélisation par homologie de la structure de l'enzyme^{5,6}. Ceux-ci utilisent des algorithmes de modélisation différents et des bases de données indépendantes.

Nous avons fourni à SWISS MODEL la structure de tbPHGDH (3DDN) comme template et celui-ci a généré un modèle de PHGDH sous sa forme tétramérique. En effet, ce webserveur est capable de générer l'état d'oligomérisation de l'enzyme si celui-ci est présent dans le template. Le modèle obtenu est présenté à la Figure 4-3. Nous observons sur ce celui-ci un tétramère asymétrique comme dans notre template. Nous obtenons un *Global Model Quality Estimation* (GMQE) de 0,69. Cette valeur est comprise entre 0 et 1 et plus elle est élevée, plus le modèle est considéré comme fiable^{81,82}. Le GMQE représente une estimation de la qualité du modèle qui combine les propriétés de l'alignement template-cible et la recherche du modèle. De plus il intègre le QMEAN, une fonction de score qui estime la qualité globale et locale du modèle en se basant sur différents propriétés géométriques⁸³.

Le second webserveur, i-TASSER, recherche des templates en se basant sur la similarité de repliement (ou de super structure) afin de former un modèle complet de l'enzyme en combinant les templates. Pour la modélisation de la structure de l'enzyme PHGDH, le programme a utilisé tbPHGDH (1YGY) comme unique modèle, renforçant par la même occasion, notre hypothèse de départ : la forme de l'enzyme bactérienne de *Mycobacterium tuberculosis* est le meilleur template à notre disposition. Remarquons

⁵ <http://prodata.swmed.edu/promals3d/promals3d.php>

⁶ <https://zhanglab.ccmb.med.umich.edu/>

que les structures de tbPHGDH 1YGY et 3DDN sont différenciées par leurs ligands mais ne présentent pas de différences structurales. Le modèle obtenu est représenté à la Figure 4-3. i-TASSER indique que le modèle a un C-score de 1,93. Ce score permet d'évaluer la qualité du modèle généré. Il se base sur la concordance entre les templates et le modèle et sur les paramètres de convergence de la simulation d'assemblage de la structure. Il est compris entre -5 et 2, plus cette valeur est élevée et plus le modèle est fiable. i-TASSER fournit un modèle pour le monomère uniquement.

Les deux modèles obtenus ont été alignés avec la structure de la forme tronquée de PHGDH afin de vérifier que ceux-ci sont en accord avec les informations structurales connues de la forme humaine de l'enzyme (Figure 4-3). Un RMSD de 2,3 Å et de 2,3 Å a été calculé pour l'alignement de la structure disponible (2G76) avec le modèle de SWISS MODEL et le modèle de i-TASSER respectivement (calculés sur bases des carbones alphas des acides aminés communs aux deux structures). L'alignement des deux modèles générés présentent un RMSD de 1,1 Å. Nous voyons donc que les deux webserveurs ont fournis des modèles très semblables et en accord avec la structure cristallographique existante. Nous avons décidé de conserver le modèle généré par SWISS MODEL car celui-ci permet d'avoir des informations sur l'état d'oligomérisation de l'enzyme.

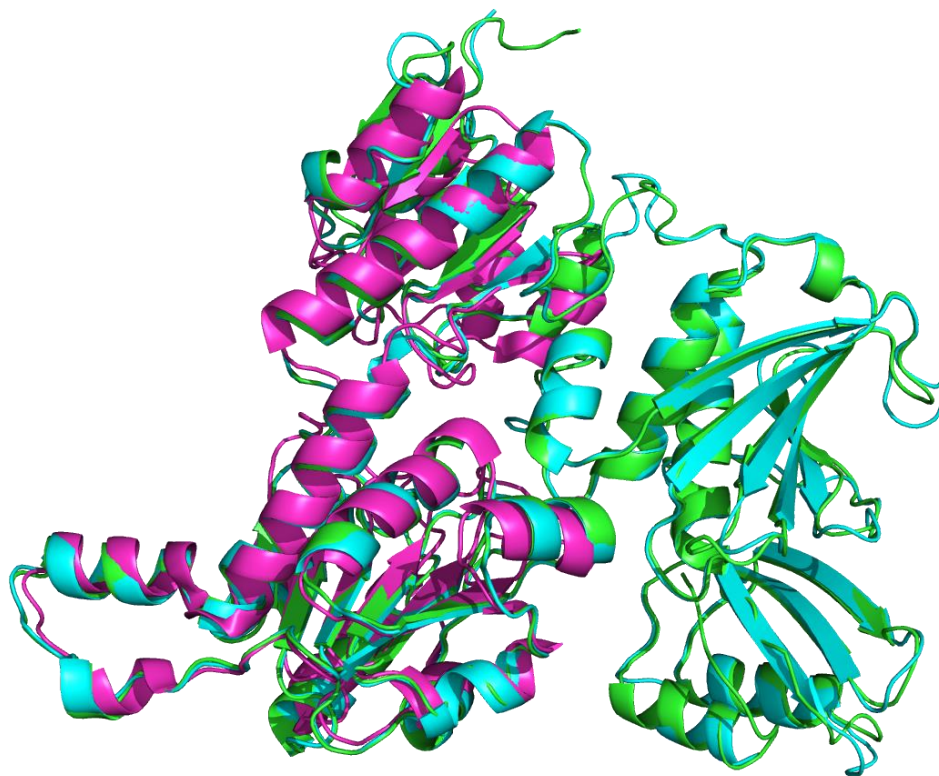


Figure 4-3: Alignement structurale de la structure cristallographique de sPHGDH (2G76 en rose), du modèle généré par SWISS MODEL (en cyan) et du modèle généré par i-TASSER (en vert).

La Figure 4-3 montre que le modèle intègre parfaitement les informations structurales disponibles. Nous avons décidé de remplacer les acides aminés modélisés par leurs équivalents provenant de la structure cristallographique quand ceux-ci sont disponibles. De cette façon, nous avons généré un nouveau modèle « hybride » comprenant des informations expérimentales et calculatoires. Ce modèle sous forme tétramérique est présenté à la Figure 4-4. De cette façon, le modèle intègre aussi les coordonnées du cofacteur et du ligand présent dans la structure de sPHGDH (2G76) : le NAD⁺ dans la poche du cofacteur et le malate dans la poche du ligand.

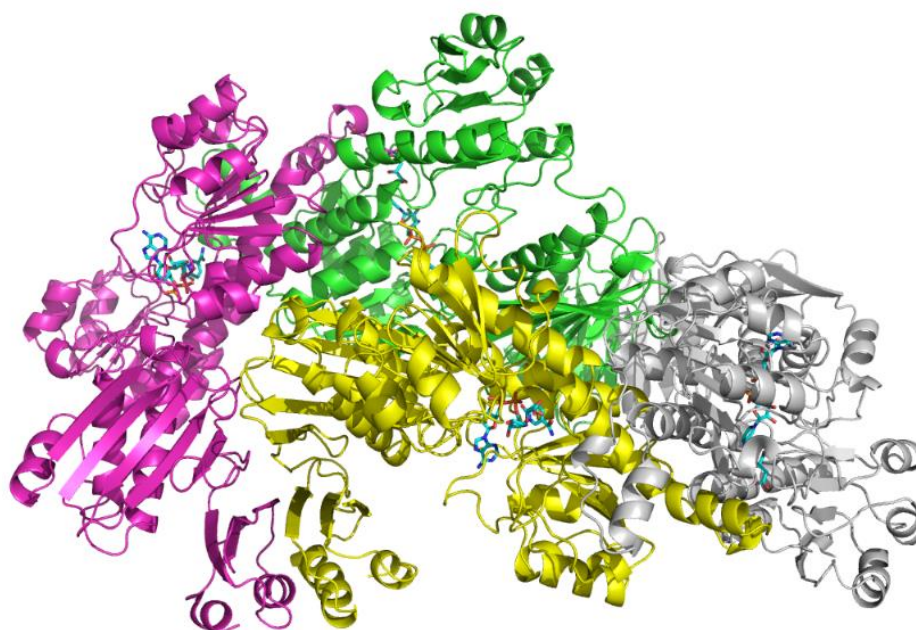


Figure 4-4: Modèle hybride de l'enzyme PHGDH complète sous sa forme tétramérique (chaque monomère est représenté dans une couleur différente).

4.2 Validation expérimentale du modèle obtenu

Nous avons obtenu dans la section précédente un modèle hybride de PHGDH sous forme d'un tétramère. Les deux webserveurs utilisés indiquent un modèle fiable. Afin de valider expérimentalement celui-ci, nous avons décidé d'employer une méthode de plus en plus utilisée pour caractériser les interactions protéines-protéines : la spectroscopie de masse couplée au cross-linking (XL-MS)⁸⁴.

4.2.1 Principe et conception expérimentale

La spectroscopie de masse couplée au cross-linking consiste à lier de façon covalente les segments des enzymes en contact. Dans notre cas, nous souhaiterions lier les 4 monomères de PHGDH formant le tétramère. Le nouveau complexe formé est ensuite digéré à l'aide d'une ou plusieurs protéases. Un enrichissement des peptides peut ensuite être réalisé avant d'analyser l'échantillon en spectrométrie de masse (Figure 4-5).

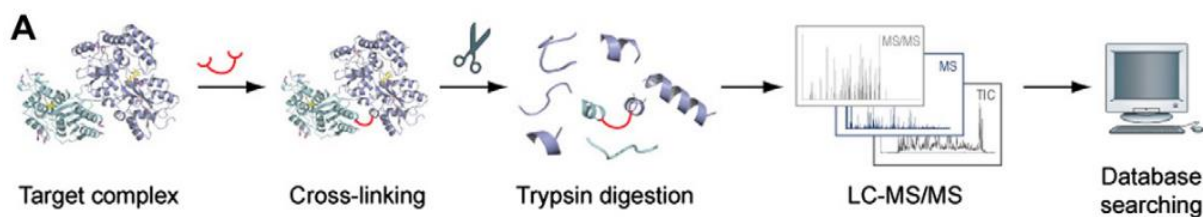


Figure 4-5: Schématisation du couplage du processus de cross-linking et de la spectrométrie de masse⁸⁵.

Lors du processus présenté à la Figure 4-5, plusieurs paramètres sont importants à prendre en compte et peuvent grandement influencer les résultats.

Le premier est le choix du *cross-linker*. Celui-ci est défini par le nombre de carbones liant les deux groupements réactifs (la « distance d'espacement »). Il a été montré qu'un cross-linker avec une distance d'espacement de 8 Å peut lier deux peptides jusqu'à une distance de 14 Å alors qu'une molécule d'espacement de 12 Å permet d'atteindre des distances jusqu'à 30 Å. Dans le premier cas, les peptides obtenus permettront d'identifier des zones de contact proches de la protéine au détriment du nombre de couplages observés. A l'inverse, l'utilisation d'un *cross-linker* avec une plus grande distance d'espacement permet d'obtenir un nombre de couplages plus important mais pas d'informations structurales précises. Il a toutefois été montré par *Leitner et al.* que l'utilisation d'un *cross-linker* plus petit ne permettait pas toujours d'obtenir des informations plus précises, mais que la quantité de peptides observés était beaucoup plus faible (deux fois moins de *cross-links* observés)⁸⁶. Nous avons décidé d'utiliser comme *cross-linker* le BS3 (Bissulfosuccinimidyl suberate) représenté à la Figure 4-6. Cette molécule présente une distance d'espacement de ~ 11 Å. Elle est composée de huit atomes de carbones et peut lier deux lysines séparés par une distance pouvant atteindre jusqu'à 30 Å. Le choix de ce composé a été motivé par sa solubilité dans l'eau, et par sa longue distance d'espacement permettant d'augmenter les chances d'observer des peptides liés. De plus, le fait que nous ayons déjà utilisé ce *cross-linker* dans le cadre d'une collaboration, nous savons qu'il permet de lier de façon covalente le tétramère de PHGDH⁶⁹. La concentration de *cross-linker* utilisée lors du couplage est également importante. En effet, à faibles concentrations, peu de peptides seront liés et il ne sera possible d'identifier que peu ou pas de zones de contact. A l'inverse, l'utilisation de

concentrations trop importantes peut favoriser la formation de liaisons aspécifiques ne provenant pas du tétramère de PHGDH et rendant l'interprétation des résultats impossible. Nous avons donc décidé de travailler à 0,5 et 5mM en BS3.

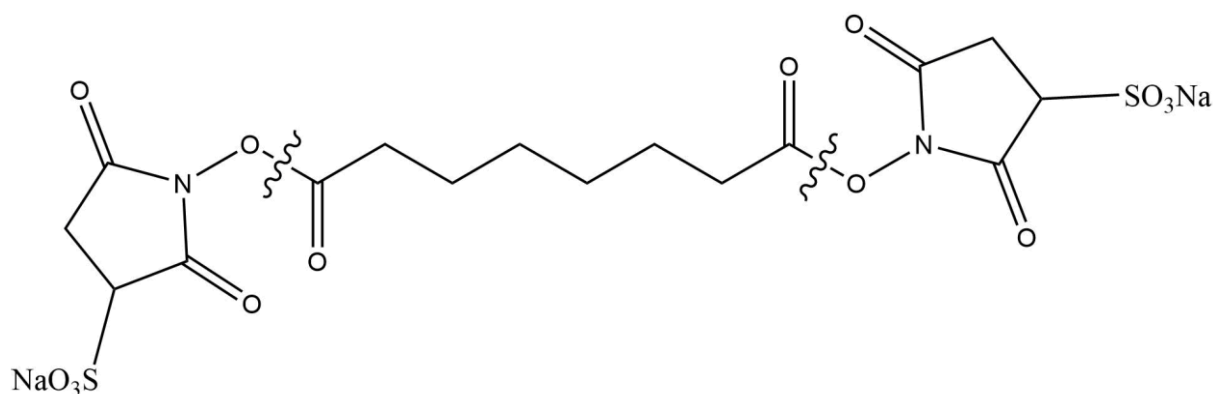


Figure 4-6: Structure du BS3 (Bissulfosuccinimidyl suberate), la liaison brisée lors de la réaction de cross-linking est mise en évidence.

Finalement, il est important de choisir des enzymes de digestion permettant d'avoir des peptides de tailles raisonnables. En effet, si la taille des peptides est trop grande, ceux-ci peuvent être liés plusieurs fois et former des « *high order cross-linked peptides* ». Leur formation rend l'identification de ces peptides impossible lors de l'analyse des résultats. Nous avons par conséquent décidé d'utiliser plusieurs enzymes de digestion, seules ou en mélange, afin de garantir l'obtention de peptides analysables : lysC ; chymotrypsine ; trypsine ; chymotrypsine/trypsine.

4.2.2 Résultats obtenus

Les échantillons ont été analysés par LC-MS-MS (nanoRSLC Ultimate 3000, Thermofischer ; maXis Impact Q-TOF, Bruker) et les résultats interprétés avec le programme Stavrox⁸⁷. Le Tableau 4-1 reprend le nombre de peptides liés par *cross-linking* en fonction des conditions expérimentales. Afin d'identifier ces peptides, Stavrox a déterminé les peptides possibles en fonction de l'enzyme de digestion ou du

mix utilisé. Il a ensuite cherché les combinaisons possibles pour la liaison de deux peptides avec une masse supplémentaire de 138,07Da (correspondant au BS3 après réaction). Finalement, nous prenons en compte les peptides présentant un FDR (False Discovery Rate) inférieur à 0,05, afin d'éliminer les combinaisons ayant de fortes chances d'être des faux positifs. Un FDR inférieur à cette valeur correspond à moins d'un faux positif tous les 2000 peptides. Le FDR est une valeur statistique qui a été mise en place pour les manipulations conventionnelles de spectrométrie de masse. Il tient compte de la qualité des données, de l'identification des fragments et de la possibilité d'identification de faux positifs (en lien avec une base de données). Dans le cadre spécifique du *cross-linking*, nous avons une base de données réduite imposant de choisir un FDR plus stringant. Nous avons dès lors choisi de ne considérer que les peptides liés présentant un FDR inférieur à 0,05. Nous ferons toutefois la distinction entre les peptides liés de très haute qualité présentant un FDR inférieur à 0,01 et ceux de bonne qualité dont le FRD est compris entre 0,01 et 0,05. Un exemple de résultat obtenu pour la condition trypsine-5mM BS3 est présenté à titre illustratif à l'Annexe A.

Tableau 4-1: Tableau récapitulatif du nombre de combinaisons de peptides uniques identifiés comme étant "cross-liqué" en fonction des conditions de cross-linking et de digestion.

PROTÉASE(S)	[BS3]	NOMBRE DE PEPTIDES UNIQUES LIÉS
Trypsine	5mM	30
Chymotrypsine + trypsine		22
LysC		0
Trypsine	0,5mM	9
Chymotrypsine		0
Chymotrypsine + trypsine		2

Sur le Tableau 4-1, nous voyons clairement que la concentration en BS3 influence directement le nombre de peptides liés. Nous observons aussi que la trypsine est la seule

enzyme de digestion permettant d'observer de nouveaux peptides après couplage. Cette peptidase digère la liaison peptidique de la protéine au niveau des lysines et des arginines. Les mesures de spectrométrie de masse étant réalisées en milieu acide, les acides aminés basiques comme la lysine ou l'arginine sont donc chargés. La charge en début ou fin du peptide facilite l'identification de celui-ci. La LysC ne coupe quant à elle que les lysines, ce qui entraîne la formation de plus grands peptides. Ces peptides sont plus susceptibles de sortir de la gamme d'analyse (m/z doit être compris entre 300 et 1600 Da et une limitation à 5 charges). De plus, la fragmentation de ces peptides est aussi plus complexe, ce qui rend leur séquençage par le programme plus difficile. Finalement, la chymotrypsine digère l'enzyme après une méthionine, isoleucine, sérine, thréonine, valine, histidine, glycine ou une alanine. En conséquence, les peptides générés ont plus de risque de se trouver sous la gamme d'analyse de l'appareil. De plus, la chymotrypsine est moins spécifique que les deux autres protéases employées, ce qui entraîne la formation de peptides qui ne seront pas reconnus par le programme d'analyse. Finalement, l'ionisation des peptides formés est moins bonne, entraînant une moins bonne détection de ceux-ci. Pour conclure, il est donc tout à fait normal d'observer un nombre plus élevé de peptides liés avec la trypsine dans le cadre d'une manipulation de *cross-linking* couplé à la spectrométrie de masse.

Deux visualisations des résultats obtenus sont présentées. Sur la première, nous pouvons voir la séquence primaire de PHGDH non colorée quand nous n'avons pas observé de peptides ayant subi un *cross-linking*, en vert quand nous avons observé un peptide lié correspondant à notre modèle, en orange quand cette séquence présente un peptide lié dans les limites possibles du BS3 et en rouge quand le peptide lié n'est pas compatible avec notre modèle (Figure 4-7). La Figure 4-7B représente une vue structurale de ces mêmes résultats. Les 4 monomères formant le tétramère sont représentés avec un dégradé de bleu (chaque bleu représente un monomère différent). Les peptides liés ne correspondant pas au modèle généré sont représentés en rouge, les autres sont colorés en vert.

A

1	SMAFANLR	KVLISDSLDP	CCRKILQDGG	LQVVEK	QNL	SKEELIAELQDCEGLIVRSATKVTADVINA	AEKL	71
72	QVVGRAGTGVDNVDLEA	ATRKGILVMNTPNGNSLSAAELTCGMIMCLARQIPQATASMKDGKWERK						137
138	KFMGTE	LNKTLGILGLGRIGR	EV	ATRM	QS	FGMK	TIGYDPIISPEVSASF	206
207	TPLLPSTTGLLNDNTFA	QCKKGVRV	VNCARGGIVDEG	ALLRALQSGQCAGAA	LDVFTEEP	PRDRALVD		274
275	HENVISCPHLGASTKEA	QSR	CGEEIAVQFVDMVKGKSLTG	VVNAQALTS	AFSPHTKPWIGLAE	ALGTL		342
343	MRAWAGSPKGTIQVITQ	GTSLKNAGNCLSPAVIVG	LLKEASKQ	ADVNLVNAKLLVKEAG	LNVTTS	SHSPA		411
412	APGEQGFGECLLAVALAGAPYQ	AVGLVQGTTPVLQQLNGAVFRPEVPLRRDLPLLLFRTQTSDP	AML	P				479
480	TMIGLLAEAGVRLLSYQ	TSLSVSDGETW	HVMGISSLLPSLEAWKQHVTEAF	QFHF				534

B

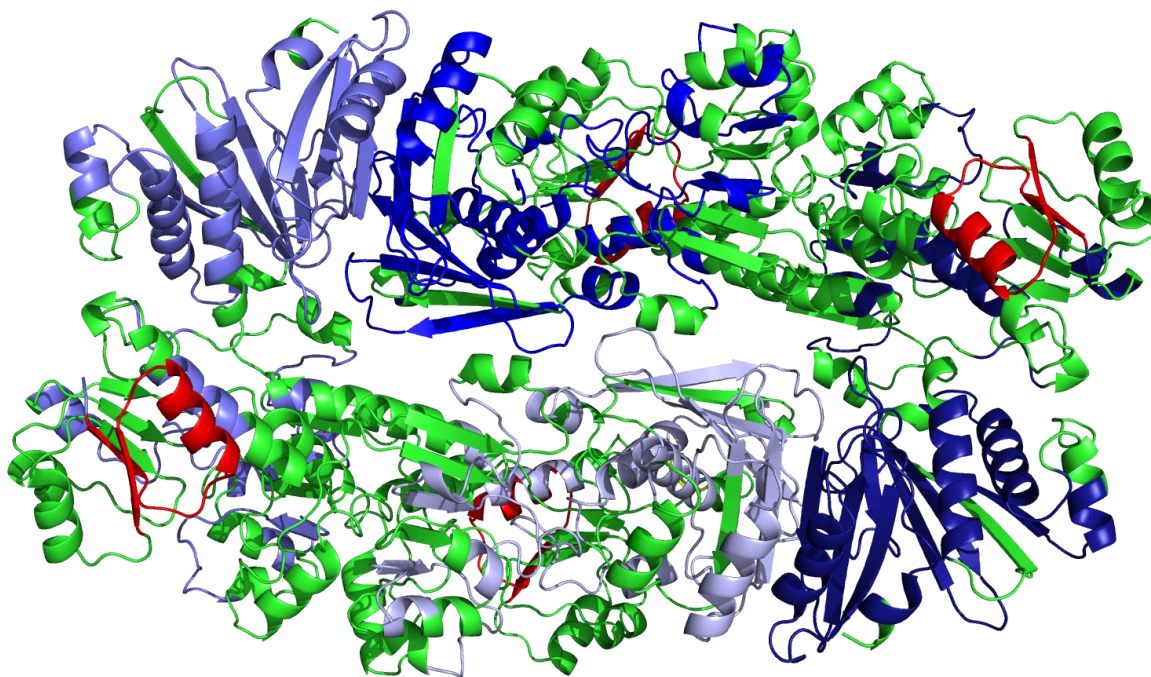


Figure 4-7: Représentation des résultats obtenus pour le cross-linking de PHGDH. (A) Séquence primaire de PHGDH (en rouge les peptides en désaccord avec le modèle, en orange à la limite de liaison de BS3 et en vert les peptides concordant avec le modèle. (B) Modèle tétramérique de PHGDH (chaque bleu représente un monomère différent, en rouge les peptides en désaccord avec le modèle, en vert les peptides en accord avec le modèle).

La Figure 4-7 montre que les peptides liés observés sont bien répartis sur la séquence. Nous observons toutefois que nous n'avons pas d'informations relatives à la région allant de l'a.a. 396 à 521. Cet écart peut être expliqué par le faible nombre de lysines dans cette région, entraînant d'une part un faible nombre de peptides potentiels entrant dans la gamme d'analyse de la machine, et d'autre part, peu de possibilités de *cross-linking*. Nous observons aussi cette répartition sur notre modèle mettant en évidence des peptides liés sur l'entièreté de la structure. Les peptides ne correspondant pas à notre modèle sont tous observés dans les 33 premiers acides aminés. Nous savons tout d'abord que ce domaine est flexible dans la structure dimérique de la forme tronquée, de plus la

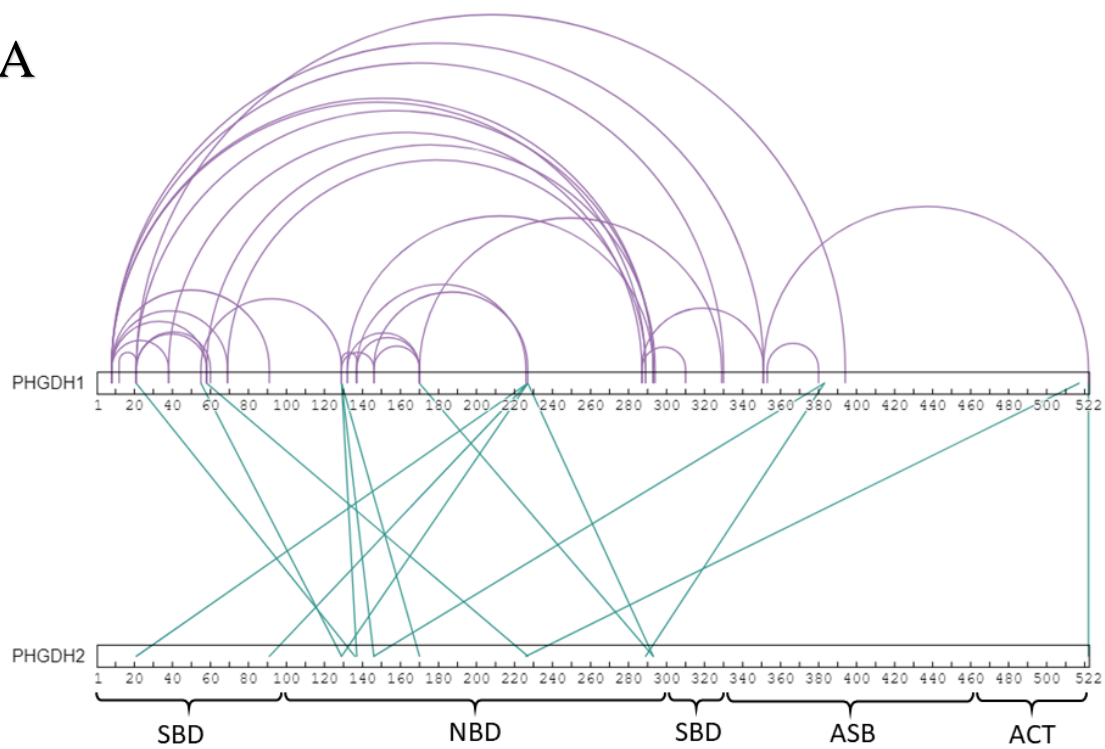
région N-terminal d'un enzyme présente une plus grande flexibilité que le cœur de celle-ci. Enfin, nous avons utilisé la structure cristallographique de cette partie de l'enzyme, il y a donc peu de chance que cette partie de modèle soit incorrecte. L'hypothèse est que suite à la flexibilité de la région N-terminal et à une concentration élevée en BS3 (tous les peptides liés ne correspondant pas au modèle ont été observés à 5mM en BS3) qui favorise les interactions aspécifiques, nous observons des peptides liés qui ne correspondent pas aux interactions en lien avec la forme tétramérique de notre enzyme.

Finalement, nous observons un total de 50 peptides liés uniques. Parmi ceux-ci, cinq ne correspondent pas à notre modèle (10%) et sont situés dans la région N-terminal de l'enzyme, cinq sont dans les limites d'interprétation du BS3 (10%) et 40 sont en accord avec le modèle (80%). Lors de l'analyse, trois peptides ont été écartés. Ceux-ci n'étaient pas en accord avec le modèle généré et présentaient un score à la limite du seuil de 0,05 fixé. De plus, l'attribution des pics de fragmentation, la séquence d'un des deux peptides ou bien la déviation par rapport à la masse théorique n'était pas bonne.

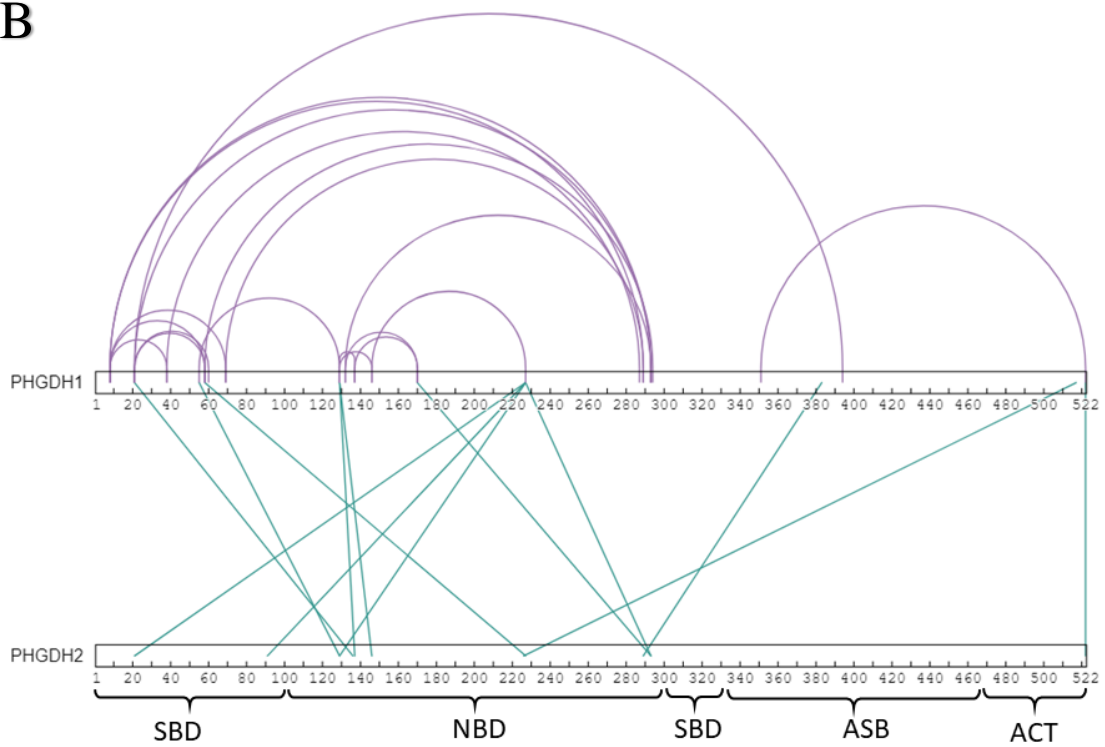
La Figure 4-7 permet de voir la couverture des peptides liés sur la séquence primaire et la structure ternaire de l'enzyme. Toutefois, ces liens se font entre deux peptides (information non représentée sur la figure précédente) et peuvent être intra-protéine et attester du repliement de l'enzyme, ou bien inter-protéine et montrer les contacts présents dans le tétramère de l'enzyme. La Figure 4-8 montre ces contacts avec une fiabilité croissante. La Figure 4-8A représente toutes les combinaisons de peptides liés uniques que nous avons obtenus et qui correspondent au modèle (nous avons exclu les 5 combinaisons incompatibles avec le modèle). Nous pouvons y voir un grand nombre de combinaisons qui permettent de montrer que le repliement du monomère généré est proche de la réalité. Le grand nombre de combinaisons inter-protéine observé indique que l'orientation des différents monomères semble correcte. La Figure 4-8B et C montre les combinaisons observées avec une plus grande astringence. Nous voyons tout d'abord (Figure 4-8B) qu'en imposant un $FDR < 0,01$, nous conservons des combinaisons sur l'entièreté du monomère, aussi bien en intra- qu'en inter-protéine. La troisième image (Figure 4-8C) ne montre que les combinaisons obtenues à 0,5mM en BS3. Cette concentration réduit les chances de couplage aspécifique et toutes les combinaisons

observées à cette concentration présentent un FDR inférieur à 0,01. Nous y voyons un nombre réduit de combinaisons. Toutefois, cette image (Figure 4-8C) nous permet de voir que la région N-terminale de l'enzyme possède des combinaisons fiables. Les interactions intra-protéines couvrent encore une bonne partie de l'enzyme bien que les interactions inter-protéines soient réduites. Sur cette image tout les peptides couplés observés correspondent au modèle. La Figure 4-8B montre un nombre croissant de *cross-link* intra- et inter-protéines. Finalement la Figure 4-8A représente l'entièreté des peptides couplés obtenus. La plupart des résultats obtenus ne correspondant pas au modèle fait partie de ce jeu de données. Nous pouvons y observer que le grand nombre de couplages intra- et inter-protéines recouvre une grande partie de la séquence.

A



B



C

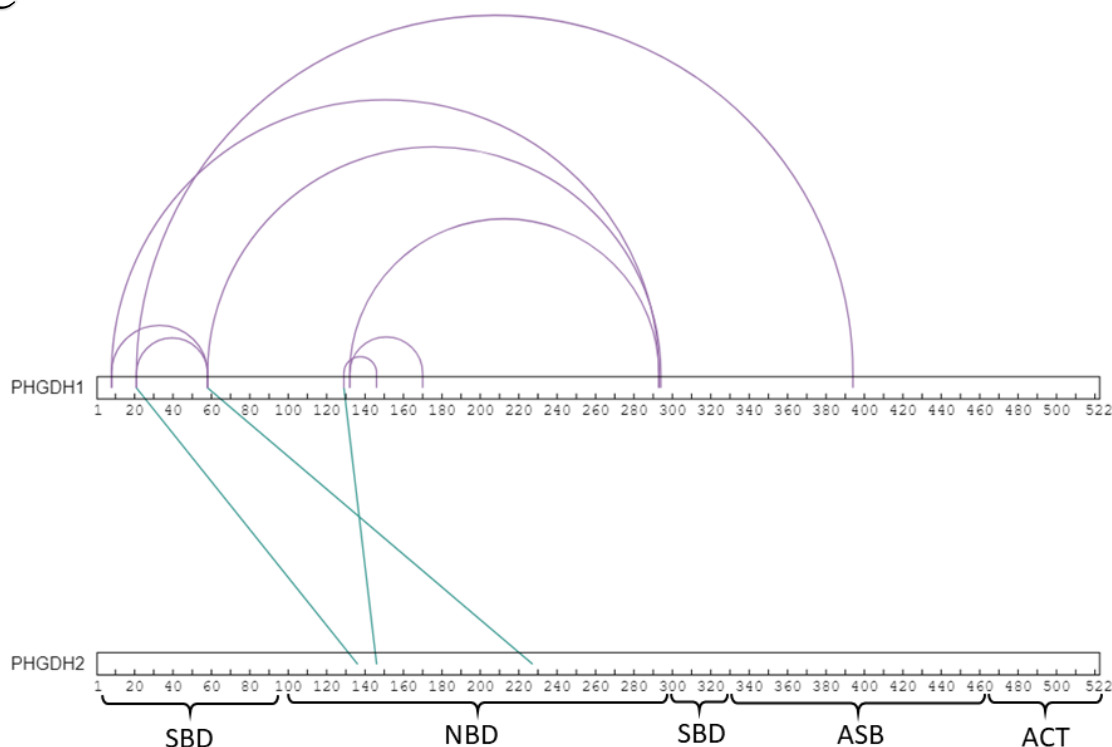


Figure 4-8 : Représentation schématique des combinaisons de peptides en accord avec le modèle. (A) Peptides liés à 5mM et 0,5mM en BS3 avec un FDR>0,05. (B) Peptides liés à 5mM et 0,5mM en BS3 avec un FDR>0,01. (C) Peptides liés à 0,5mM en BS3 avec un FDR>0,01.

Les résultats obtenus permettent de valider le repliement du monomère et les interactions inter-protéines indiquent que l'orientation des différents monomères au sein du tétramère est tout à fait plausible.

Toutefois, il est important de rappeler que le BS3 est un agent de couplage permettant de visualiser les interactions jusqu'à 30 Å. Par conséquent, nous pouvons montrer l'orientation et le repliement de PHGDH mais nous n'avons pas la précision pour montrer la position exacte d'une chaîne latérale.

Finalement, ce modèle pourrait être utilisé afin d'améliorer la compréhension de l'enzyme dans sa forme active. Il fournit également un outil pour comprendre où et comment les inhibiteurs actuellement découverts peuvent interagir.

4.3 Comparaison du modèle aux données existantes

Le modèle ici développé est un hybride composé de la structure cristallographique de l'enzyme PHGDH tronquée et du modèle généré sur base d'une forme bactérienne de l'enzyme. Nous avons donc pu conserver les ligands co-cristallisés avec la forme humaine : le malate et le NAD⁺. Tous les acides aminés de la poche liant le malate étant présents dans la forme tronquée de l'enzyme, notre modèle n'apporte pas de nouvelles informations à ce sujet. Toutefois la Figure 4-9 montre que la poche du NAD⁺ est complétée par la partie de l'enzyme qui n'avait pas été cristallisée. Nous pouvons voir d'une part la complétion de la poche catalytique et d'autre part visualiser la surface de l'enzyme montrant une poche bien définie pour la liaison du NAD⁺. Ces deux éléments expliquent la forte de liaison de PHGDH pour son cofacteur⁶⁰. Nous observons sur la Figure 4-9B, la position du NAD⁺ par rapport au site de liaison au substrat (dans ce cas-ci le 3-PG a été remplacé par le malate). Il est en effet obligatoire que le NAD⁺ soit présent dans la poche catalytique avec le substrat. Dans le cas contraire, il serait impossible au NAD⁺ d'accéder à la poche catalytique, confirmant l'hypothèse faite dans la littérature stipulant que l'ordre de liaison à la poche catalytique doit être le NAD⁺ suivi du 3-PG⁵⁵. Nous n'observons toutefois pas d'interactions entre le NAD⁺ et la partie de la poche modélisée. Le template ne contenant pas de NAD⁺, il est possible que la poche soit dans une conformation plus ouverte et que la liaison du NAD⁺ entraîne une « fermeture » de celle-ci suite aux interactions formées avec celui-ci.

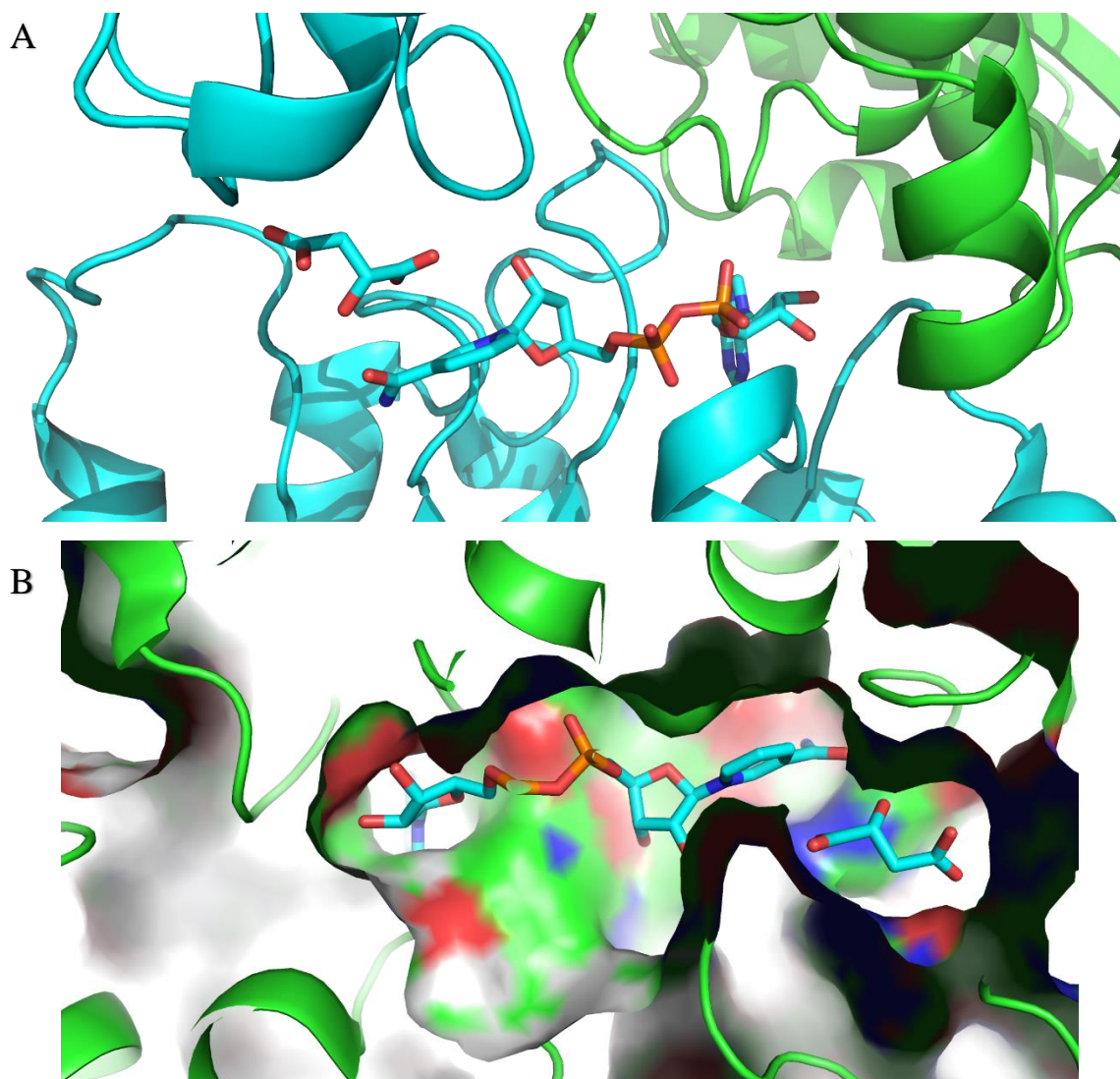


Figure 4-9: Mise en évidence de la poche de liaison de NAD⁺ dans le modèle hybride de PHGDH. (A) Position du NAD⁺ dans la poche catalytique, en cyan la structure provenant de la cristallographie (2g76) et en vert la structure générée à l'aide de méthode computationnelle. (B) Surface du modèle générée avec mise en évidence de la poche du NAD⁺ et du malate (en vert la structure du modèle, en cyan les deux ligands : le malate et le NAD⁺).

L'activité de l'enzyme complète a été comparée à celle de l'enzyme tronquée par notre collaborateur, le Dr. Edouard Mullarky (Université de Cornell, USA). L'affinité des deux enzymes pour leur cofacteur a aussi été mesurée par celui-ci. Ces résultats sont présentés à la Figure 4-10. Nous pouvons voir que l'enzyme complète présente une activité bien plus élevée que la forme tronquée et que sPHGDH bien que sous forme dimérique et tronquée conserve une légère activité. D'autre part, nous voyons que la constante de dissociation (K_D) de la forme tronquée est près de 200 fois plus élevée que celle de l'enzyme sauvage. Ces résultats montrent expérimentalement une différence majeure lors de la liaison de l'enzyme au NAD⁺ qui peut être expliquée par une poche

de liaison incomplète mise en évidence sur le modèle que nous avons généré. De plus, il est montré que cette forme tronquée est moins active que la forme sauvage, ce qui peut être dû au fait que celle-ci soit sous forme dimérique et non tétramérique, mais aussi à l'affinité réduite du cofacteur pour cette forme.

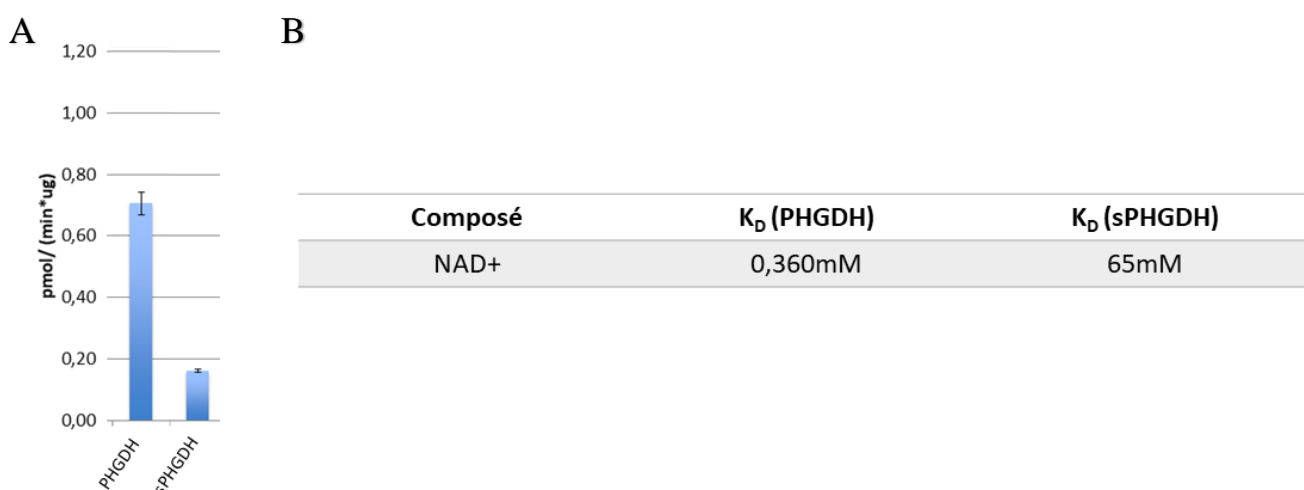


Figure 4-10: Activité spécifique et activité pour le cofacteur de l'enzyme complète et de l'enzyme tronquée (manipulation réalisée par Dr. Edouard Mullarky). (A)Activité spécifique des enzymes PHGDH complète et sPHGDH exprimée en pmol/(min* μ g). (B) Constante de dissociation du NAD⁺ mesurée par SPR.

4.4 Conclusion et perspectives

Au terme de ces travaux, nous avons montré que la forme bactérienne provenant de *Mycobacterium tuberculosis* de PHGDH est un bon *template* pour la modélisation de l'enzyme sous sa forme complète et tétramérique. Deux webserveurs indépendants (i-TASSER et SWISS MODEL) ont permis de générer ce modèle avec un haut degré de fiabilité. Les deux modèles générés présentaient une structure semblable du monomère. Une version hybride a pu être construite en conservant la structure de sPHGDH (2G76) et en y ajoutant les acides aminés manquants, générés par le webserveur.

La combinaison du *cross-linking* à la spectrométrie de masse a permis de mettre en évidence des preuves expérimentales suggérant que notre modèle est plausible. En effet, les contacts observés grâce aux peptides liés montrent que le repliement global de PHGDH et l'orientation des monomères dans la forme tétramérique de l'enzyme sont

possibles. Toutefois, bien que ces informations soient encourageantes, elles ne permettent pas de confirmer que notre modèle correspond à la structure exacte de l'enzyme. D'une part, l'interprétation des données dans le cas d'un tétramère asymétrique sont délicates et peuvent être rapidement surinterprétées. D'autre part, le BS3 ayant une distance d'espacement élevée, des interactions à longue distance peuvent être observées. Afin d'obtenir des informations complémentaires, des expérimentations de dichroïsme circulaire pourraient être effectuées. Celles-ci ne permettraient pas de valider le modèle généré mais mettraient en évidence les éléments de structure secondaire de l'échantillon qui pourraient être comparés avec ceux présents sur PHGDH modélisé.

De façon plus générale, la résolution de la structure de PHGDH par d'autres méthodes est difficile. En effet, l'hétérogénéité du système (dimère-tétramère dur à séparer) et la « petite taille » de l'enzyme rendent l'obtention de la structure de celle-ci par cryo-microscopie électronique difficile. Au contraire, cette taille est considérée comme grande pour une résolution par résonance magnétique nucléaire. Une résolution de la structure par cette méthode avec une protéine de cette taille serait longue et fastidieuse.

Finalement, nous avons mis en évidence la modélisation de la poche du NAD⁺ montrant clairement que celle-ci est incomplète dans la structure de PHGDH tronquée. Ces observations sont en accord avec l'ordre de liaison proposé dans la littérature et avec les expérimentations réalisées par notre collaborateur montrant que l'enzyme complète a une meilleure affinité pour le cofacteur.

5. Etude du site actif de PHGDH

5 Etude du site actif de PHGDH

5.1 Prédiction de la conformation du substrat au sein du site actif

Afin de mieux comprendre la liaison du substrat au sein du site actif de PHGDH nous avons décidé de prédire sa position au sein de celui-ci. Le *docking* du 3-phosphoglycérate a été réalisé sur le modèle précédemment généré. Le site actif étant compris dans le domaine cristallisé, nous connaissons les coordonnées du ligand co-cristallisé dans le site du substrat : le malate. Ceci nous a permis de valider la méthode.

5.1.1 Validation de la méthode de docking

La protéine et le malate ont été préparés avec le logiciel Discovery Studio visualizer 2016 (minimisation de l'énergie par application d'un champ de force) et la simulation de la position du ligand au sein du site actif a été réalisée avec le logiciel GOLD. Un exemple d'input est présenté à l'Annexe B. La fonction de score CHEMPLP a été utilisée pour donner un score à chaque conformation de la molécule dans le site actif. La génération de la molécule est faite dans une sphère de 10 Å centrée sur le centre de masse du ligand présent dans la structure cristallographique (malate). 20 conformations ont été générées pour chaque molécule. La Figure 5-1 représente la position du malate cristallisé en superposition avec la position calculée au sein du site actif.

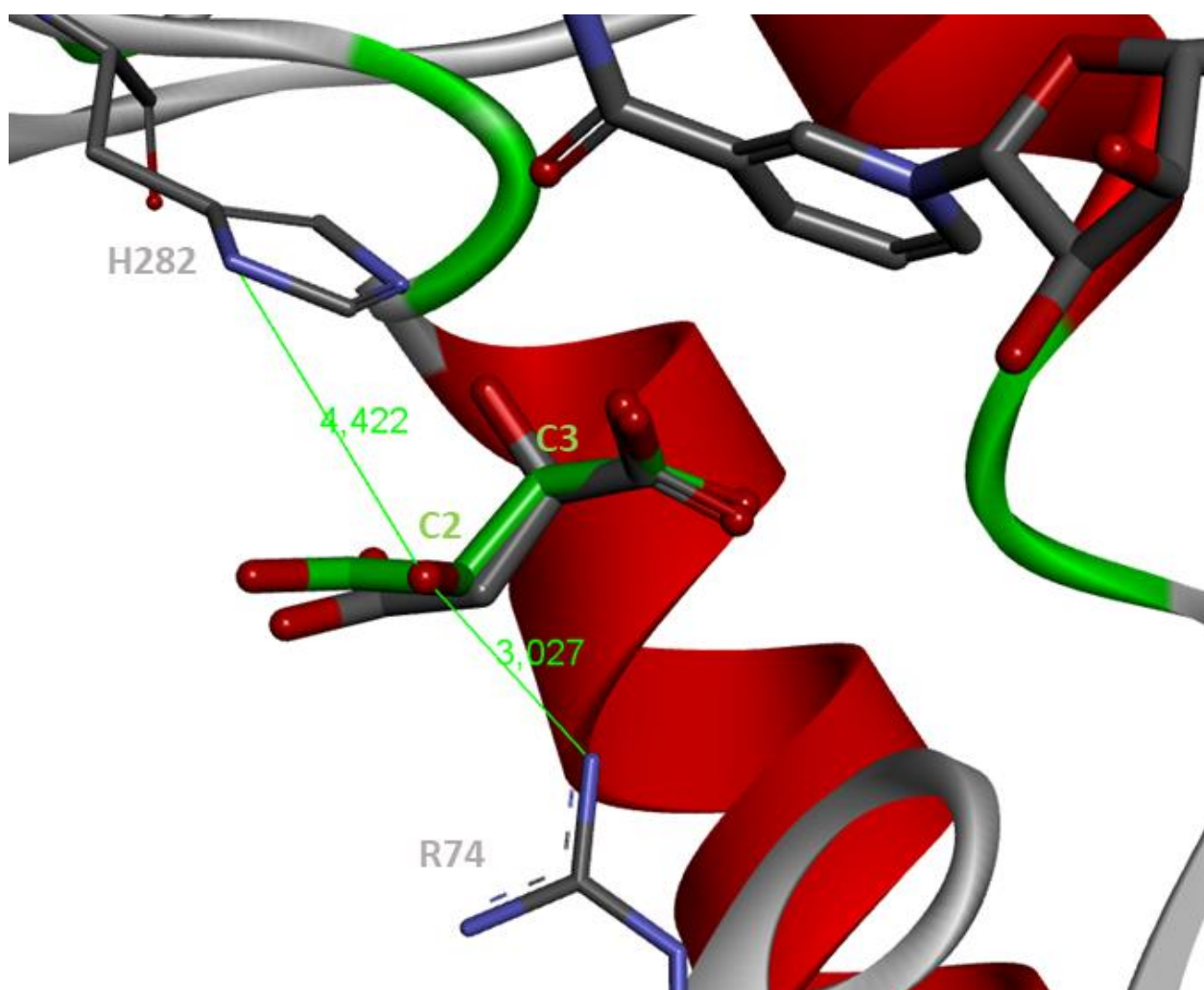


Figure 5-1: Conformation du malate obtenu par docking (en vert) et cristallisé (2G76, en gris) dans le site actif du modèle de PHGDH.

La position du ligand au sein du site actif ainsi que son mode de liaison sont bien conservés. Nous observons que les positions du C2 et C3 ont été inversées. La stabilisation de l'alcool (C2) par la formation d'un pont H avec l'H282 observée pour la conformation cristallisée présente un angle (O-H...N) moins favorable dans la conformation générée par docking. Ce changement est compensé par une interaction supplémentaire par pont hydrogène entre cette même fonction alcool et l'R74. Il est probable que cette interaction supplémentaire ait induit cette position différente de l'alcool. La position du malate au sein de la poche étant conservée et le mode de liaison globalement reproduit, nous considérons que notre méthode est validée et peut-être utilisée pour la génération du mode de liaison du 3-PG au sein du site actif.

5.1.2 Liaison du substrat au sein de la poche catalytique

Maintenant que nous savons que notre méthode est validée, nous pouvons l'utiliser pour des ligands dont le mode de liaison n'est pas connu. Le substrat naturel de l'enzyme, le 3-phosphoglycérate, dont le mode de liaison dans une PHGDH de type 1 est inconnu, a donc été testé. Les interactions formées au sein du site actif sont présentées à la Figure 5-2, Figure 5-3 et Tableau 5-1.

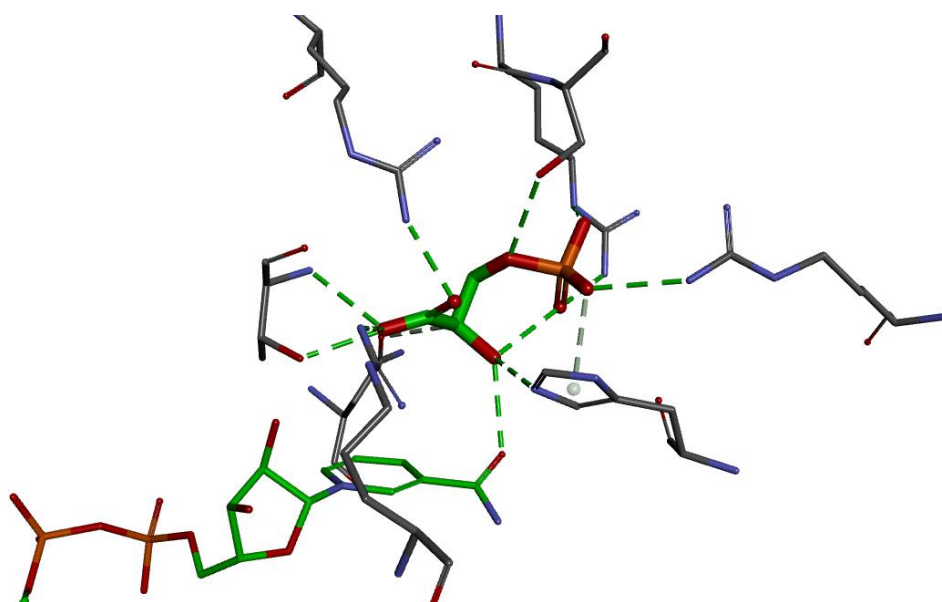


Figure 5-2: Interactions formées par le 3-phosphoglycérate au sein du site actif de PHGDH

Tableau 5-1: Interactions par pont hydrogène formées entre le site actif et son substrat.

Fonction 3-pg	Acide aminé	Distance (D-H...A) (Å)	Angle (D-H...A) (°)
Phosphate	R53	2,6	136
Phosphate	R53	2,9	92
Phosphate	S54	3	92
Carboxylate	R74	2,6	125
Carboxylate	T77	2,9	137
Phosphate	R134b	2,9	139
Carboxylate	R235	3,0	155
Alcool	H282	3,0	155

Nous y voyons que le groupement phosphate est stabilisé par les acides aminés R53, S54 et H282 ainsi que par R134 provenant d'un autre monomère (Figure 5-2). Un second site d'ancrage est observé au niveau du carboxylate et est stabilisé par des interactions avec R74, T77 et R235. Finalement, la fonction alcool, qui doit être oxydée lors de la réaction de l'enzyme, est stabilisé par l'H282. Ce mode de liaison montre que la fonction alcool qui doit être oxydée est orientée vers le NAD⁺. Etant donné que le phosphohydroxypyruvate, le produit de réaction de PHGDH en présence de 3-phosphoglycérate, a été co-cristallisé dans la forme de PHGDH provenant de l'organisme *Mycobacterium tuberculosis*, nous avons aligné la position du 3-PG avec celui-ci au sein du site actif de l'enzyme (Figure 5-3). Le mode de liaison de ces deux composés au sein du site actif de la forme humaine de PHGDH est représenté à la Figure 5-3B.

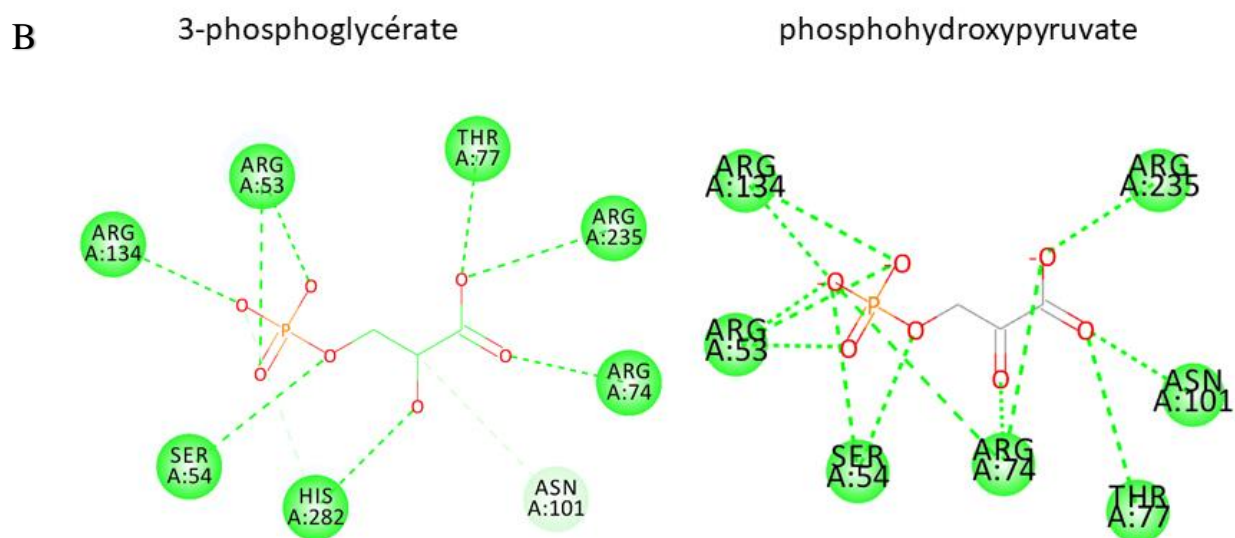
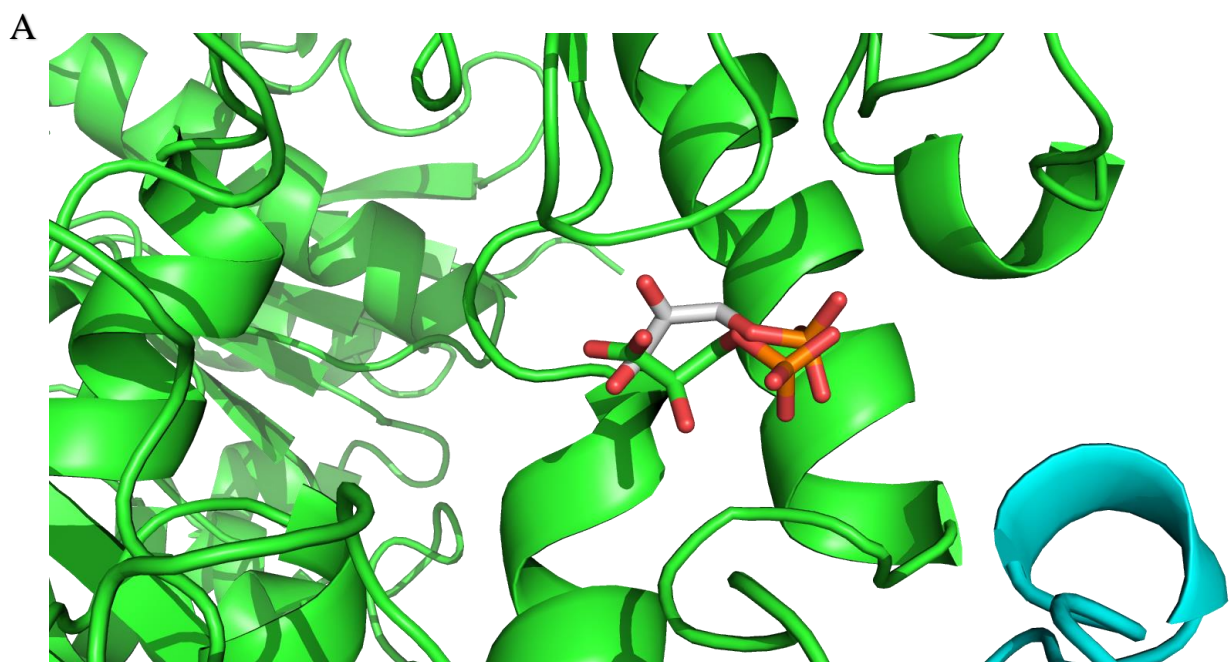


Figure 5-3: Comparaison de la liaison au sein du site actif de hPHGDH du 3-phosphoglycérate généré et de l'hydroxypyruvate (3DDN). (A) Superposition de la structure du 3-PG et Ppyr au sein du site actif de hPHGDH (deux monomères de PHGDH sont représentés en vert et cyan ; le 3-PG est en verte et le Ppyr est en gris).

Nous observons sur la Figure 5-3 que la position du 3-PG au sein du site actif est très proche de celle du Ppyr. De plus, le mode de liaison est fortement conservé. Nous retrouvons la stabilisation du phosphate par R53, S54 et R134B. La fonction carboxylate est toujours en interaction avec R74, T77, et R235. La différence majeure se situe dans la position du groupement carbonyle réactif. Cette différence de position peut être expliquée par le fait que le phosphohydroxypyruvate n'a pas été cristallisé en présence

de NAD⁺. De plus, afin de libérer le produit après réaction, il est logique que celui-ci ait une moins bonne affinité pour l'enzyme que le substrat. Ces changements peuvent expliquer la modification de position de l'oxygène au sein de la poche.

Cette comparaison entre la conformation du 3-phosphoglycérate que nous avons générée et celle du phosphohydroxypyruvate qui a été co-cristallisé dans tbPHGDH a permis de mettre en évidence les deux positions de stabilisation du substrat au sein de la poche.

5.2 Interactions observées au sein du site actif de hPHGDH

5.2.1 Comparaison des modes de liaison de ligands connus

Dans la partie précédente, nous avons mis en évidence deux points d'ancrage nécessaires à la bonne liaison du substrat et du produit au sein de la poche du site actif. Afin de confirmer l'importance de ces régions du site actif, nous avons comparé le mode de liaison de 3-phosphoglycérate avec celui des ligands connus se liant au sein de l'enzyme. La Figure 5-4 représente la superposition du 3-PG, du tartrate (5N6C) et du malate (2G76) liés au site actif de PHGDH. Nous observons que les deux ligands connus sont stabilisés par les mêmes points d'ancrage que ceux que nous avons observés pour le 3-PG. Il est aussi intéressant de noter que le malate favorise la stabilisation de l'alcool par le NAD⁺ et H282 comme observé pour le 3-PG. Ces observations confortent les résultats obtenus par docking.

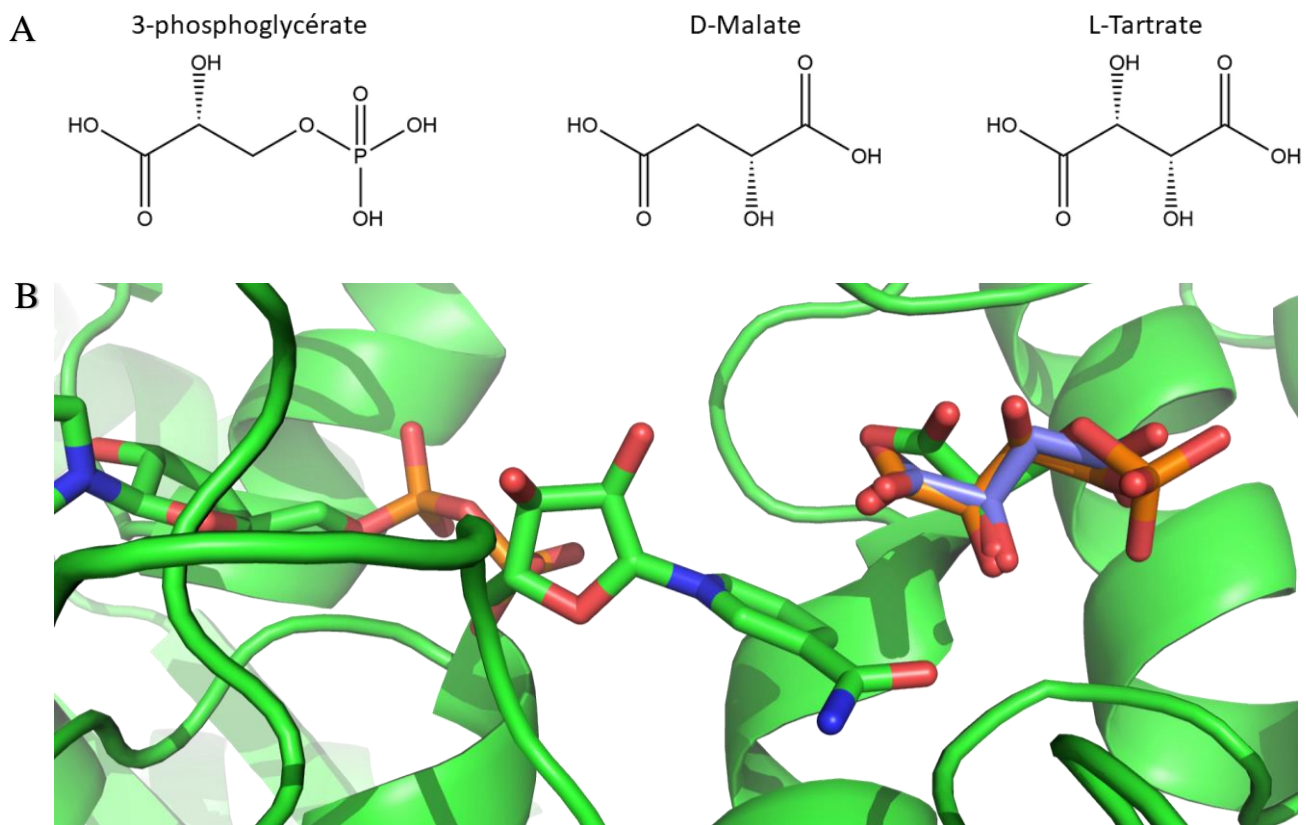


Figure 5-4: Structure et conformation du 3-phosphoglycérate, du tartrate et du malate au sein du site actif de PHGDH. (A) Structure des 3 molécules. (B) Conformation du 3-phosphoglycérate (docking, en vert), du malate (2G76, en bleu) et du tartrate (5N6C, en orange)

5.2.2 Conservation des acides aminés interagissant avec le 3-phosphoglycérate.

Les différentes formes de PHGDH sont classées dans 3 catégories : type 1, type 2 et type 3 (H ou K). Nous avons aligné les séquences de plusieurs enzymes de chaque type de PHGDH afin d'observer la conservation des acides aminés impliqués dans la liaison des différents ligands (Tableau 5-2).

Tableau 5-2: Récapitulatif des acides aminés impliqués dans la liaison des ligands du site actif et leur conservation dans les différents types de PHGDH. NA= non applicable.

ACIDE AMINÉ	3-PG	MALATE (2G76)	TARTRATE (5N6C)	TYPE 1	TYPE 2	TYPE 3H	TYPE 3K
R235	x	x	x	Conservé	Conservé	Conservé	Conservé
R53	x	x		Conservé	Conservé	Conservé	Conservé
S54	x	x	x	Conservé	Conservé	Conservé	Conservé
R74	x		x	Conservé	Non conservé	Conservé	Conservé
T77	x	x	x	Conservé mammifère	Non conservé	Non conservé	Non conservé
R134b⁷	x	x	x	Conservé	Non conservé	Non conservé	NA
H282	x	x	x	Conservé	Conservé	Conservé	Non conservé
Asn101	x			Conservé	Conservé	Non conservé	Non conservé

Le Tableau 5-2 montre que la plupart des acides aminés impliqués dans la liaison du 3-PG et des ligands du site actif de hPHGDH sont conservés dans les différentes souches de PHGDH. Il est intéressant de noter que les acides aminés T77 et R134B sont spécifiques aux PHGDH de Type I. De plus, des résultats préliminaires (réduction de l'entropie de surface, voir chapitre 1, section 3.3) ont montré que le mutant 1 (mutation de la séquence ER134KK en AR134AA) est complètement inactif. Sur base de ces informations, il est intéressant de vérifier l'importance de l'arginine 134 dans la liaison du 3-PG. Cet acide aminé a la particularité d'être le seul provenant d'un autre monomère de PHGDH et d'être spécifique aux enzymes de type 1.

⁷ "b" signifie que l'acide aminé provient du second monomère.

5.3 Arginine 134 : Etude structurale et fonctionnelle

5.3.1 Effet de la mutation R134A sur l'activité de l'enzyme

La forme mutée de l'enzyme comprenant la mutation R134A a été produite, exprimée et purifiée de la même façon que la forme non mutée de l'enzyme (voir chapitre 1). Une enzyme pure a été obtenue et utilisée pour déterminer l'activité de l'enzyme. La Figure 5-5 représente l'évolution de la vitesse initiale en fonction de la quantité de substrat pour l'enzyme mutée et sa forme sauvage. Le Tableau 5-3 reprend les valeurs de K_m et V_{max} extraites de la Figure 5-5.

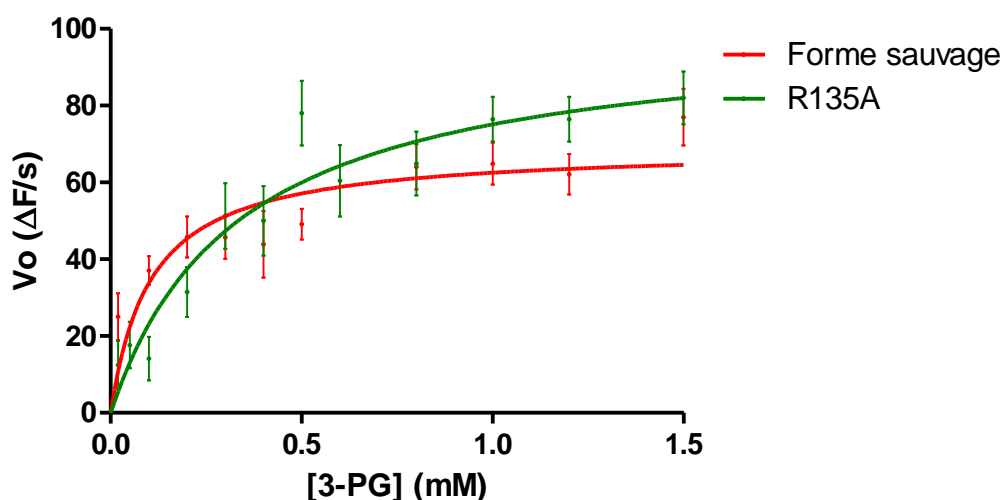


Figure 5-5: Vitesse initiale en fonction de la concentration en 3-phosphoglycérate de l'enzyme sauvage (en rouge) et de sa forme mutée R134A (en vert).

Tableau 5-3: Constante cinétique de PHGDH sous sa forme sauvage et comportant la mutation R134A.

	FORME SAUVAGE	R134A
Km (mM)	0,107 ± 0,032	0,337 ± 0,098
Vmax (ΔF/s)	69,35 ± 4,94	100,4 ± 10,22

La Figure 5-5 et le Tableau 5-3 montrent une perte d'affinité du mutant R134A pour le substrat. En effet, le Km du mutant R134A est de 0,337 ce qui est plus de trois fois supérieur à celui de l'enzyme sauvage. Nous observons aussi une augmentation de Vmax dans le cadre du mutant. Nous expliquons ce phénomène par le fait qu'à basse concentration, une baisse d'affinité se traduise par une perte d'activité mais qu'en augmentant la concentration du substrat, on favorise son insertion dans le site actif. A haute concentration, une plus basse affinité pour le substrat au travers d'une mutation entraîne une baisse d'affinité pour le produit (l'arginine stabilisait le substrat et le produit par interaction avec le groupement phosphate qui n'est pas modifié) et permet donc un relargage plus rapide de celui-ci. Cette hypothèse permet d'expliquer l'évolution de Vo de l'enzyme mutée par rapport à l'enzyme sauvage. Il est aussi intéressant de noter que la concentration physiologique du 3-PG dans une cellule humaine se situe entre 50 et 80μM⁸⁸⁻⁹⁰. A ces concentrations, l'enzyme sauvage est plus active que le mutant R134A, ce qui explique pourquoi cette mutation n'a pas été privilégiée par l'évolution.

Ces résultats mettent en évidence que l'arginine 134 est bien impliquée dans la liaison du 3-PG et que sa mutation réduit l'affinité de l'enzyme pour son substrat.

5.3.2 Effet de la mutation R134A sur la structure de l'enzyme et la liaison au substrat

Afin d'observer les modifications structurales apportées par la mutation R134A, nous avons inséré cette mutation dans la forme tronquée des PHGDH (sPHGDH). Le mutant a été produit, exprimé, purifié et cristallisé de la même façon que pour la forme sauvage présenté dans le chapitre 1. Un des cristaux a été obtenu dans la condition 0,1 M Bis-tris (pH=6,5), 0,2 M Li₂SO₄, 25% PEG3350 à une concentration de 15 mg/mL et un ratio protéine : réservoir de 1 : 2. Les paramètres de la collecte sont repris au Tableau 5-4. Nous remarquons une faible complétude. En effet, nous avons dû utiliser le programme STARANISO⁸ qui permet de mettre en place un *cut-off* dans la direction de la plus haute résolution. Par conséquent la complétude de notre collecte n'est pas représentative du celle de notre jeu de données.

⁸ <http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>

Tableau 5-4: Données collectées et statistiques de l'affinement (déterminée avec Molprobit), les valeurs entre parenthèses sont celles à la plus haute résolution.

SPHGDH R134A - NAD ⁺	
DONNÉES DU CRISTAL	
Groupe d'espace	P1
Dimensions	
a , b , c (Å)	43,24 ; 59,73 ; 64,84
α , β , γ (°)	88,82 ; 90,11 ; 80,13
Résolution (Å)	43,1 - 1,8 (1,9 - 1,8)
Nombre de réflexions uniques	44095 (2223)
Complétude (%)	79,89 (40,15)
I/ σ I	7,4 (3,1)
Redondance	10,3 (2,8)
STATISTIQUES	DE
L'AFFINEMENT	
Ligand	NAD ⁺
Rwork / Rfree (%)	0,21 / 0,23
Nombre d'atomes	4539
Macromolécules	4407
N° molécules d'eau	132
RMS (liaison) (Å)	0,01
RMS (angle) (°)	1,62
Ramachandran favorable (%)	95,78

Nous observons logiquement sur la Figure 5-6 l'absence de densité électronique pour le mutant R134A là où l'arginine est normalement présente dans l'enzyme sauvage. Nous notons aussi que la mutation n'a pas impacté la structure de sPHGDH. Seul un très léger déplacement (RMSD= 0,6Å) est observé entre la forme sauvage et la forme mutée. La perte d'activité ne peut donc pas être attribuée à une modification structurale de l'enzyme mais bien à une perte d'affinité pour le substrat au travers de l'interaction avec le phosphate du 3-PG.

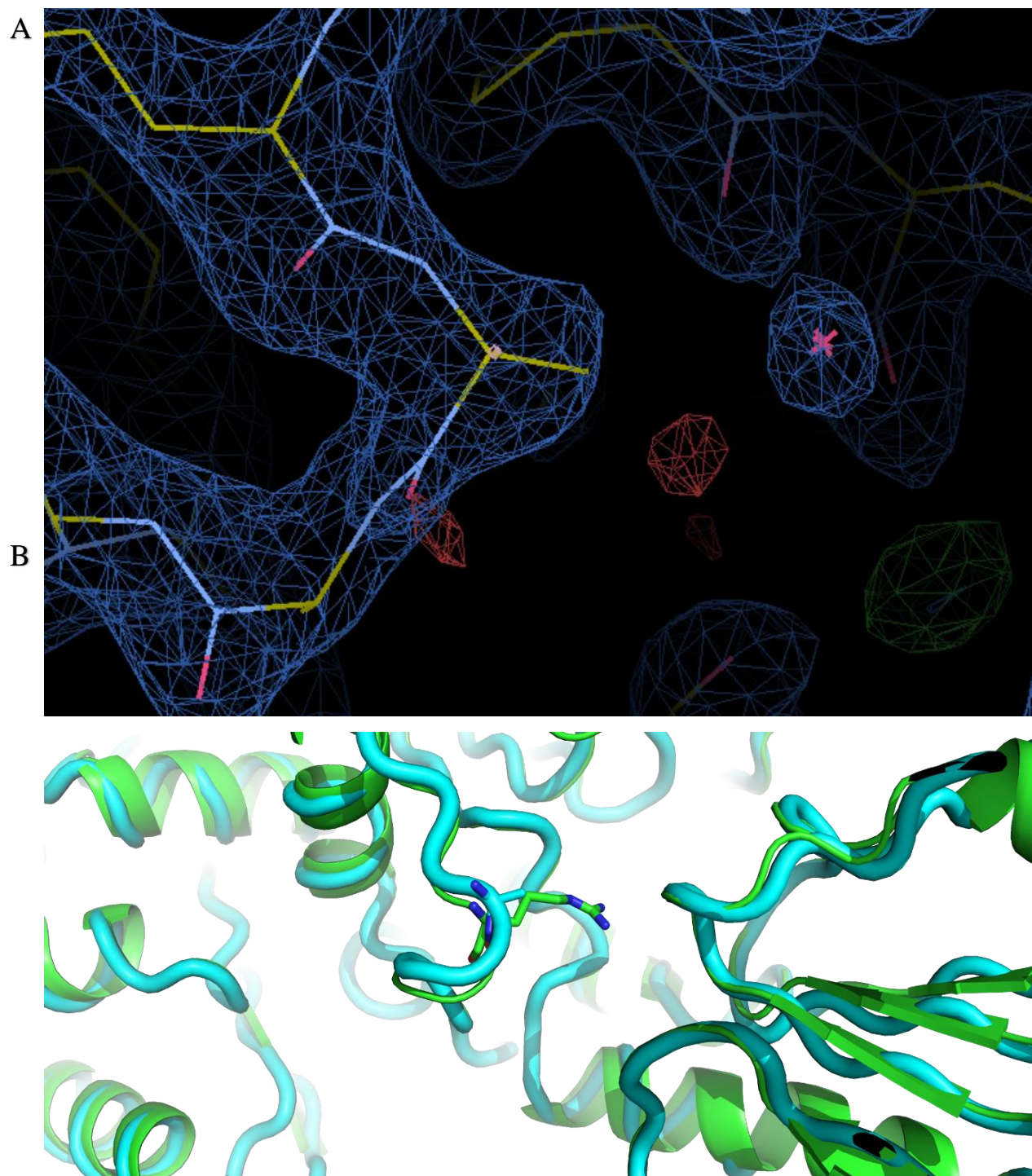


Figure 5-6: Structure de sPHGDH R134A, zoom sur la mutation. (A) Carte de densité comprenant les coordonnées après affinement. (B) Alignement de la forme sauvage de sPHGDH (2G76 en vert) et sPHGDH R134A (en cyan) (RMSD= 0,8Å).

5.3.3 Rôle de la boucle comportant l'arginine 134

Nous avons observé lors de la production des mutations en surface de la protéine (chapitre 1, section 3.3) que la mutation ERKK → ARAA rendait l'enzyme complètement inactive. Or, nous savons que la mutation R134A n'est pas suffisante pour rendre l'enzyme complètement inactive. De plus, les 3 acides aminés mutés n'entrent pas en interaction avec l'autre monomère de PHGDH et ne semblent pas affecter l'état d'oligomérisation de l'enzyme. L'effet de cette mutation sur l'enzyme n'est par conséquent pas encore clair pour nous. Afin de simuler ce changement d'interface au sein de l'enzyme, nous avons conçu 3 peptides qui pourrait venir déstabiliser l'interface entre les deux monomères : ERKK (133-136); WERKKFMGTE (132-141); MKDGKWERKKFMGTELNGKT (128-147). Ces 3 peptides ont été testés par notre collaborateur le Dr. E. Mullarky et les résultats obtenus sont présentés à la Figure 5-7.

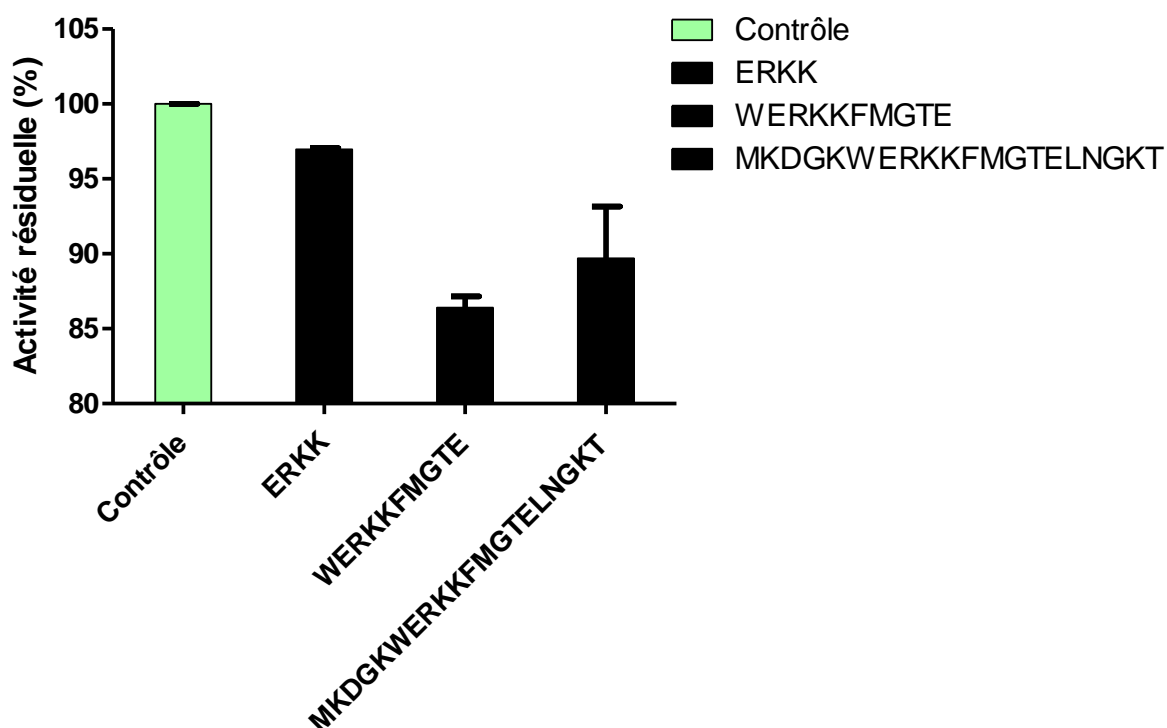


Figure 5-7: Activité résiduelle de hPHGDH en présence des 3 peptides conçus représentés dans une gamme de 80 à 100%.

La Figure 5-7 montre que le peptide « WERKKFMGTE » et le peptide « MKDGKWERKKFMGTELNGKT » induisent une légère inhibition ($86,4 \pm 1,1\%$ et $89,7 \pm 4,9\%$ d'activité résiduelle respectivement). Bien que celle-ci soit faible, elle montre que ces peptides sont capables d'interagir avec la protéine et que cette interface peut être intéressante dans le cadre de la conception d'inhibiteurs ciblant PHGDH.

5.4 Conclusion

Au terme de ce chapitre, nous avons mis en avant de façon théorique le mode de liaison de 3-phosphoglycérate au sein du site actif de l'enzyme PHGDH. Le mode de liaison proposé est concordant avec les données disponibles pour l'homologue bactérien de l'enzyme provenant de *Mycobacterium tuberculosis*.

Nous avons mis en évidence que le 3-PG se lie grâce à deux sites d'ancrage. Le premier, au niveau de la fonction carboxylate, est en interaction avec R74, T77 et R234. Le second permet la stabilisation du phosphate grâce à R53, S54, R134b. La fonction alcool est positionnée en direction du NAD⁺ grâce à la fonction amide de celui-ci et à l'H282. Nous retrouvons la plupart des acides aminés interagissant avec le 3-PG dans le mode de liaison de malate et du tartrate. Nous avons aussi mis en évidence que seuls deux de ces acides aminés sont spécifiques aux enzymes de Type I.

Nous avons muté R134 en alanine et mis en évidence que celui-ci est en effet important pour la liaison du substrat. Nous avons aussi montré que cette mutation n'entraînait pas de changement structural confirmant que l'effet observé sur la perte d'affinité du substrat est uniquement dû à la fonction de l'arginine.

Le mutant ERKK→ARAA a montré que la boucle présentait, elle aussi, une importance pour l'activité de l'enzyme. En effet une perte complète de l'activité de l'enzyme est observée après mutation de ces 3 acides aminés. Toutefois nous ne comprenons pas encore comment cette boucle est liée à l'activité de l'enzyme.

Nous avons finalement testé la capacité d'inhibition de trois peptides semblables à la boucle et mis en évidence un effet de ceux-ci sur l'activité de l'enzyme. Bien que cet

effet soit léger, il montre que l'interface de la boucle avec le site actif peut être intéressante d'un point de vue thérapeutique.

6. Criblage de composés au sein du site actif

6 Criblage de composés au sein du site actif

6.1 Génération des ensembles de molécules

Nous avons vu dans le chapitre précédent que l'utilisation du *docking* a permis de mettre en évidence le mode de liaison probable du 3-phosphoglycérate au sein de la poche catalytique. Nous avons décidé d'utiliser cette méthode afin de trouver de nouveaux ligands se liant dans la poche de liaison de 3-PG et du NAD⁺. Pour ce faire, nous avons utilisé la base de données ZINC⁹¹ pour générer deux ensembles de molécules. ZINC contient plus de 230 millions de molécules commercialement disponibles sous forme de fichiers .mol2 prêts à être utilisés en docking. Le premier ensemble formé est nommé « *drug* » et contient 2863 molécules approuvées par la FDA. Le second ensemble est constitué de 98179 composés biogéniques semblables à des médicaments et est nommé « *druglike* ». Ces composés présentent tous un poids moléculaire compris entre 150 et 500 g/mol, un logP < 5, un nombre de donneur d'hydrogène inférieur ou égal à 5 et d'accepteur d'hydrogène inférieur ou égal 10. Ces composés répondent donc à la règle des 5 de Lipinski. Ces deux bases de données ont permis d'avoir un grand nombre de composés structuralement différents et ont l'avantage de présenter des propriétés intéressantes dans des perspectives thérapeutiques.

Pour chaque ensemble nous avons généré la conformation des molécules au sein du site de liaison du cofacteur et du substrat avec la même méthode que celle utilisée pour générer la conformation du 3-PG. Les positions obtenues ont été classées en fonction de leur score (calculé à l'aide de la fonction de score Chemscore). Ce classement sera divisé en deux sections. La première présentera les molécules avec la meilleure affinité pour l'enzyme. Nous y retrouverons logiquement des plus grandes molécules car celles-ci sont capables de faire un plus grand nombre d'interactions. La seconde section normalisera le score en fonction du nombre d'atomes lourds de la molécule afin de mettre en évidence l'efficacité du ligand. De cette façon, il est possible d'identifier des ligands plus petits et plus fonctionnalisables.

6.2 Poche de liaison au cofacteur

6.2.1 Résultats obtenus pour l'ensemble « drug »

Les 5 molécules présentant le meilleur score sont présentées dans le Tableau 6-1 avec leur score associé. La Figure 6-1 montre la conformation de ces 5 composés dans le site actif en comparaison avec celle du NAD⁺.

Tableau 6-1: Les 5 composés de l'ensemble "drug" dont la conformation générée présente le meilleur score et le score associé.

N°	MOLÉCULE	SCORE
1	Nonoxymol-9	108
2	Montelukast	105
3	Fulvestrant	104
4	Fosinopril	103
5	Demecarium	101

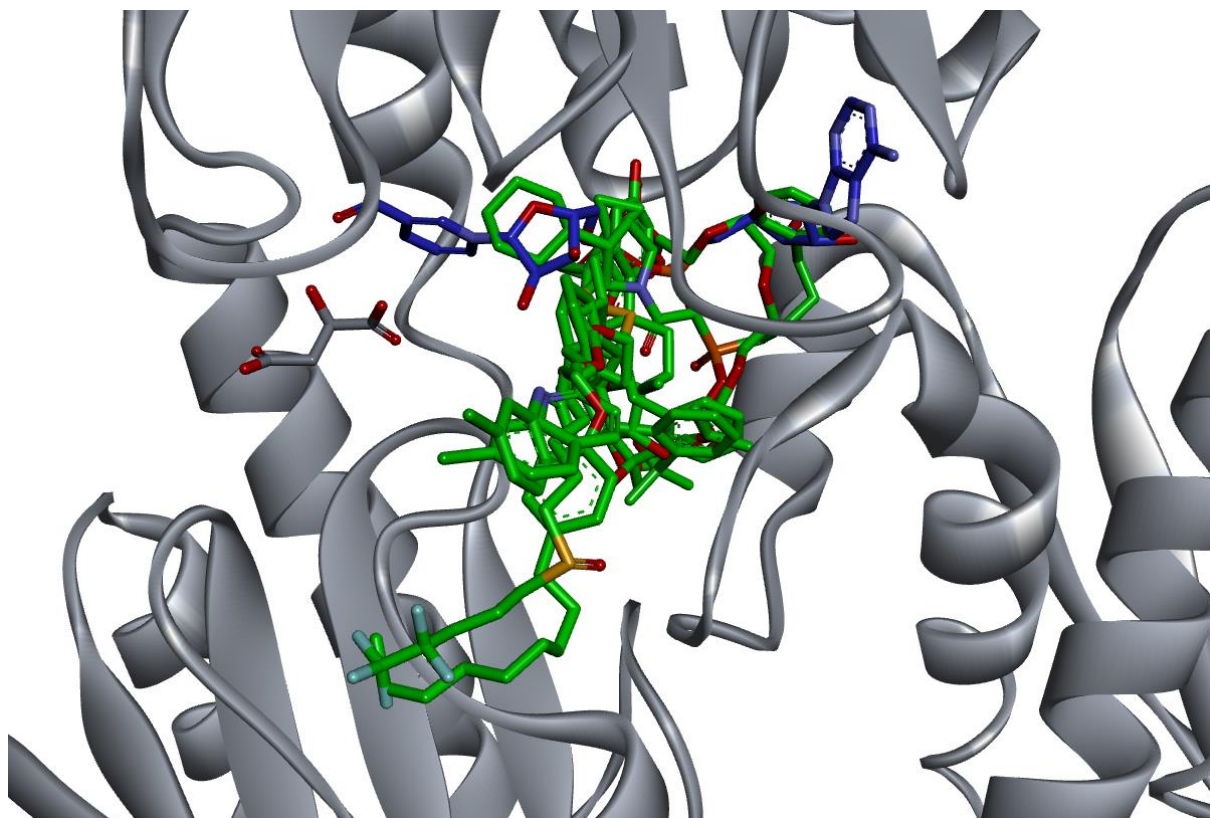


Figure 6-1: La conformation des 5 composés (en vert) obtenant le meilleur score par docking au sein de la poche du NAD⁺ (en gris) est comparée à la conformation du NAD⁺ (en bleu).

La conformation de chaque composé ainsi que son mode de liaison au sein du site actif sont présentés à l'Annexe C. Nous observons sur la Figure 6-1 que les composés ne privilégient pas des interactions au sein de la poche du NAD⁺. Le nonxymol-9, fulvestrant, le montelukast et le demecarium se positionnent à l'interface du NAD Binding Domain (NBD), du Substrat Binding Domain (SBD) et du domaine modélisé (voir chapitre 2). Les interactions formées de cette façon sont principalement de type van der Waals. Nous remarquons que toutes les molécules forment des interactions dans une partie de la poche formée par les acides aminés R154, I155, G156 et E399.

Une fois le score obtenu normalisé en fonction du nombre d'atomes lourds, nous obtenons cinq nouvelles molécules, plus petites que les précédentes. Celles-ci sont reprises dans le Tableau 6-2 et la comparaison de leur conformation avec celle du NAD⁺ est présentée à la Figure 6-2. La conformation de chaque composé et leur mode de liaison au sein du site actif sont présentés à l'Annexe C.

Tableau 6-2: Les cinq composés de l'ensemble "drug" dont la conformation générée présente le meilleur score normalisé.

N°	MOLÉCULE	SCORE NORMALISÉ
6	Hydrochlorure de guanidine	6,8
7	Glycine	6,6
8	Hydrochlorure de cystéamine	6,1
9	Oxybate de sodium	6,0
10	L-alanine	5,7

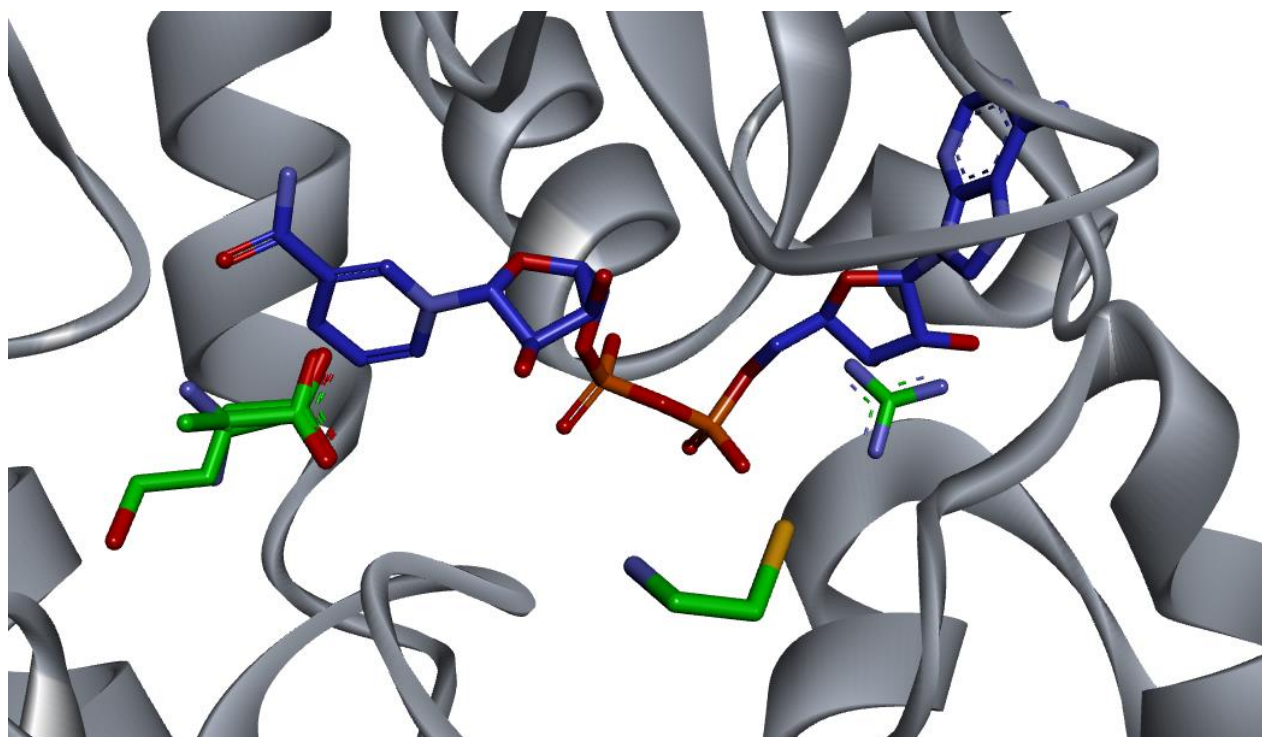


Figure 6-2: La conformation des 5 composés (en vert) obtenant le meilleur score normalisé par docking au sein de la poche du NAD⁺ (en gris) est comparée à la conformation du NAD⁺ (en bleu).

Nous observons sur la Figure 6-2, qu'aucune des molécules ne se place au sein du site de liaison du NAD⁺. L'hydrochlorure de cystéamine se place hors du site actif. Nous retrouvons l'hydrochlorure de guanidine proche de la poche de liaison au phosphate en formant une interaction par pont hydrogène avec E399. Finalement la glycine, l'alanine et l'oxybate de sodium se placent dans la poche de liaison au malate.

6.2.2 Résultats pour l'ensemble « druglike »

La conformation des composés de l'ensemble « druglike » a ensuite été simulée au sein de la poche de liaison du NAD⁺. Les composés présentant les meilleurs scores sont présentés au Tableau 6-3 et la conformation de ceux-ci comparée à celle du NAD⁺ est présentée à la Figure 6-3.

Tableau 6-3: Les cinq composés de l'ensemble "druglike" dont la conformation générée présente le meilleur score et leur score associé.

N°	MOLÉCULE	ZINC ID	SCORE
11	2-[[2-[[2-[3-(4-fluorophenyl)-7-keto-5,9-dimethyl-furo[3,2-g]chromen-6-yl]acetyl]amino]acetyl]amino]	ZINC02115415	99
12	2-[[2-[[2-(7-keto-2,5,9-trimethyl-3-phenyl-furo[3,2-g]chromen-6-yl)acetyl]amino]acetyl]amino]acetate	ZINC02118203	99
13	(1S,3R,3aS,6aS)-5'-ethyl-1-[(4-hydroxyphenyl)methyl]-5-phenethyl-spiro[1,2,3a,6a-tetrahydropyrrolo[4	ZINC15969642	98
14	4-(1H-indole-2-carbonyl)-N-[3-[2-(1H-indol-3-yl)ethylamino]-3-oxo-propyl]piperazine-1-carboxamide	ZINC12898259	98
15	3-[3-(3-aminophenyl)amino-3-(2-hydroxy-1-naphthyl)-propanoyl]-4-hydroxy-1-methyl-quinolin-2-one	ZINC03846632	98

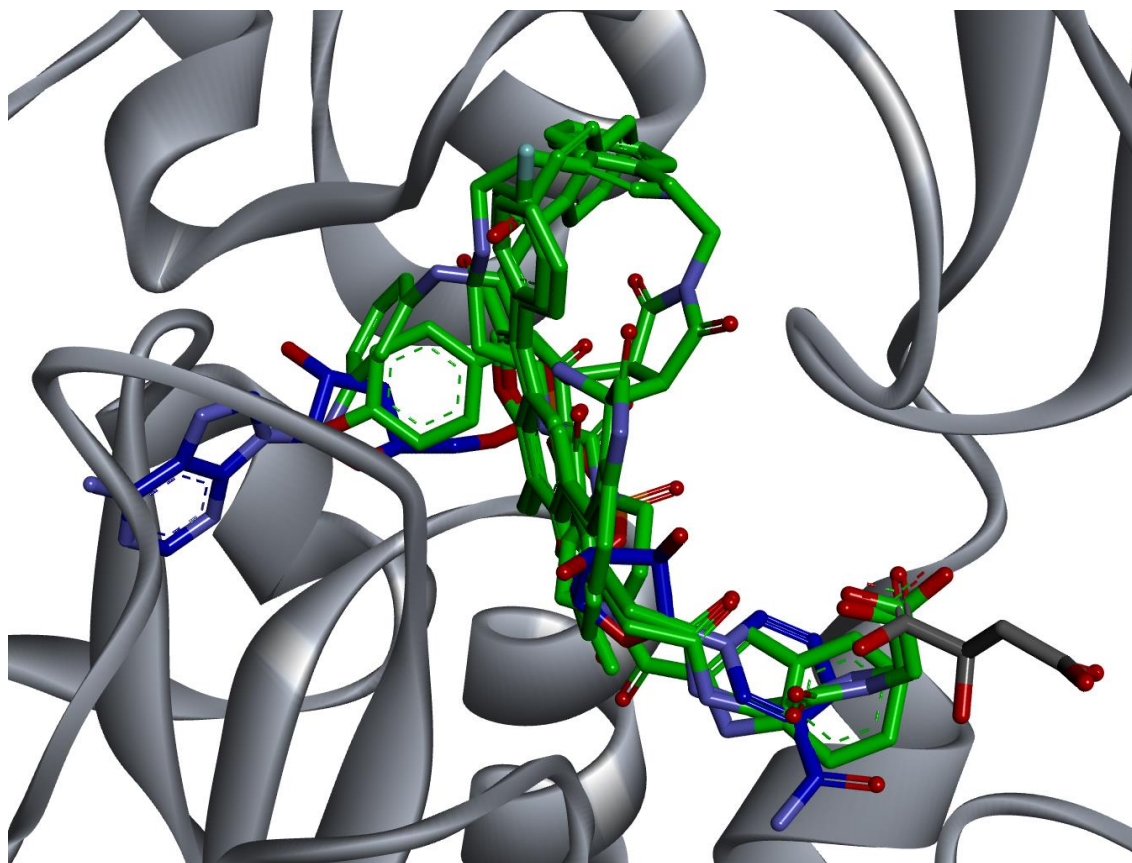


Figure 6-3: La conformation des 5 composés de l'ensemble « druglike » (en vert) obtenant le meilleur score par docking au sein de la poche du NAD⁺ (en gris) est comparée à la conformation du NAD⁺ (en bleu).

Nous pouvons voir sur la Figure 6-3 que tous les composés se positionnent au sein de la poche du NAD⁺ (le mode de liaison de chaque composé est repris à l'Annexe C). Nous remarquons aussi que les molécules **11** et **12** interagissent dans la poche du malate par pont hydrogène entre l'acide carboxylique et T77. Une poche secondaire favorisant les interactions de type π -alkyle est aussi observée et permet une stabilisation des cycles aromatiques de tous les composés. Finalement, nous observons comme précédemment que le principal point d'ancrage des molécules se situe au niveau de la partie de la poche qui lie les deux groupements phosphate du NAD⁺ grâce aux acides aminés R154, R235 et E399 principalement. Un récapitulatif de la liaison aux acides aminés du site actif pour chaque molécule est repris au Tableau 6-5.

Les scores de toutes les molécules ont ensuite été normalisés en fonction du nombre d'atomes lourds et un nouveau classement sur base des scores normalisés a été réalisé.

Celui-ci est repris au Tableau 6-4. La conformation des cinq meilleures molécules a ensuite été mise en évidence au sein du site actif de PHGDH et comparée avec celle du NAD⁺ (Figure 6-4). Le mode de liaison complet de chaque molécule est repris à l'Annexe C.

Tableau 6-4: Les cinq composés de l'ensemble "druglike" dont la conformation simulée présente le meilleur score normalisé.

N°	MOLÉCULE	SCORE NORMALISÉ
16	4-cyclohexyl-butylamine	4,4
17	D,L-buthionine	4,4
18	Acide 6-guanidinohexanoïque	4,3
19	3-phenylpropanohydrazide	4,2
20	1,2-dihydroxyindane	4,2

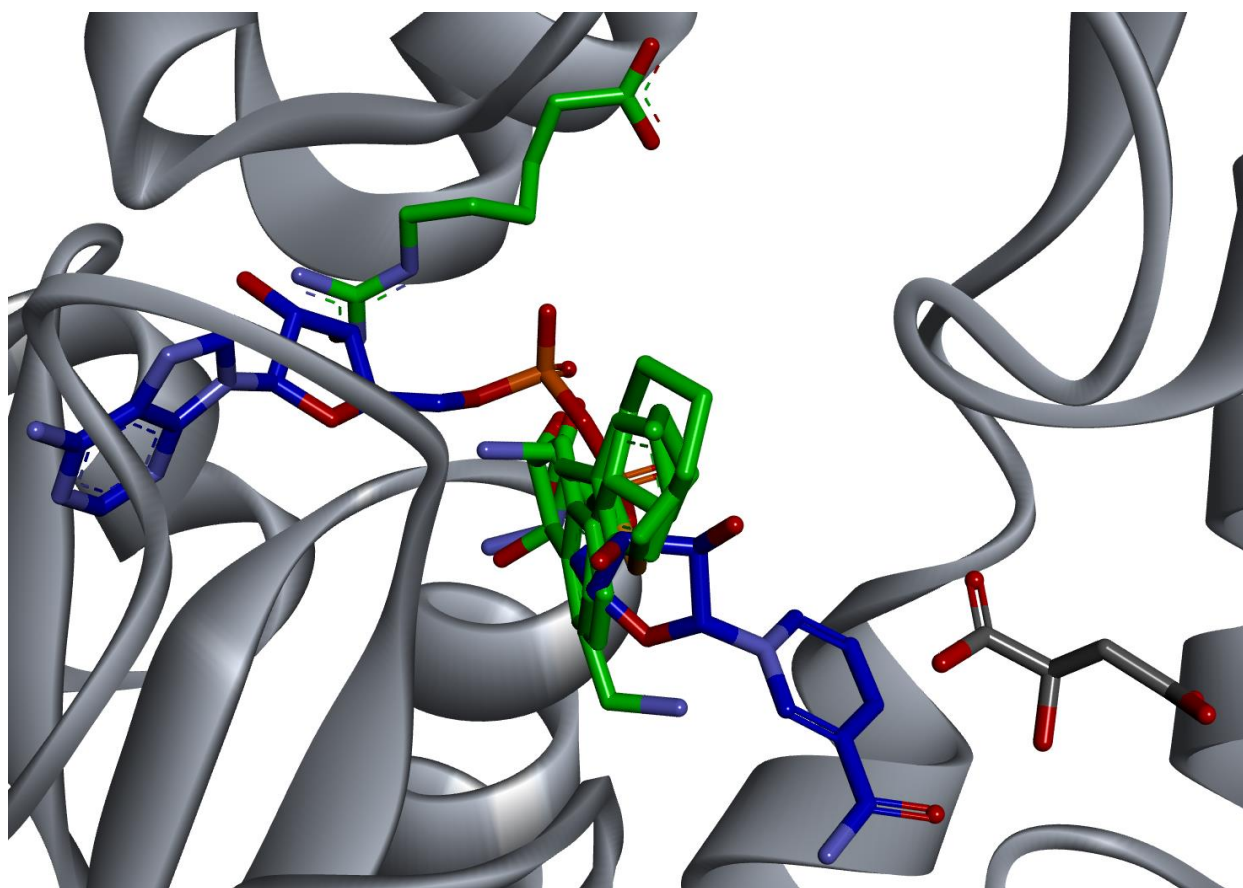


Figure 6-4: La conformation des 5 composés de l'ensemble « druglike » (en vert) obtenant le meilleur score normalisé par docking au sein de la poche du NAD⁺ (en gris) est comparée à la conformation du NAD⁺ (en bleu).

Sur la Figure 6-4 nous observons que toutes les molécules à l'exception de l'acide 6-guanidinohexanoïque se lient au centre de la poche de liaison au NAD⁺. Les molécules **17**, **19**, **20** se lient à la même position que les deux phosphates du NAD⁺ et présentent des interactions avec R154 et E399. Ces interactions sont aussi observées pour la molécule **18**, bien que ce composé ne se positionne pas dans l'espace normalement occupé par le NAD⁺. Finalement le 4-cyclohexyl-butylamine a une position plus proche du cycle tétrahydrofurane et présente une interaction par pont hydrogène avec la C233 ($d_{O...H-N} = 3\text{\AA}$, $\angle O...H-N = 7^\circ$).

Afin d'avoir un aperçu global des acides aminés entrant en interaction avec les différents composés, un tableau récapitulatif reprenant tous les acides aminés formant des interactions par pont hydrogène (en vert) ou π -cation (en orange) est représenté au Tableau 6-5. Si ces acides aminés forment aussi des interactions de type π -alkyl comme nous en observons souvent avec ces composés, celles-ci sont aussi représentées (en rose).

Nous observons dans ce tableau que les acides aminés R154, I155, E399 présentent des interactions avec le plus grand nombre de composés (≥ 5). Ceux-ci font partie de la poche de liaison des phosphates. Les acides aminés D174, H205, C233 et R235 forment chacun des interactions avec trois composés différents. Ces acides aminés font partie du pli de Rossmann qui est un motif impliqué dans la liaison du NAD⁺.

Tableau 6-5: Récapitulatif des acides aminés formant des interactions par pont hydrogène (en vert), π -cation (en jaune) ou π -alkyl (en rose) avec les composés de l'ensemble « druglike » dont la conformation a été générée dans la poche de liaison au NAD⁺.

N°	R74	T77	G78	N101	R154	I155	G156	D174	H205	C233	R235	A382	N392	E399
NAD ⁺														
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														

6.2.3 Discussion

Nous avons vu dans la section 6.2.1 que les composés provenant de l'ensemble « drug » n'ont pu se positionner dans la poche du NAD⁺. Les composés présentant le meilleur score après normalisation privilégient la poche de liaison au malate. Il est probable que le volume de la poche de liaison au NAD⁺ ne permette pas à ces composés de se lier au sein de celle-ci, privilégiant par conséquent leur liaison au niveau de la poche du substrat. Les composés présentant le meilleur score pour l'ensemble « Drug » sont tous utilisés dans divers traitements. Le nonoxymol est un surfactant possédant des propriétés spermicides. Des propriétés anti-VIH lui ont tout d'abord été attribuées⁹² avant que celles-ci ne soient réfutées⁹³. Le montelukast est quant à lui un antagoniste des récepteurs leukatriènes D4 et est donné pour le traitement de l'asthme^{94,95}. Le fulvestrant est déjà utilisé dans le cadre du cancer. Il se lie aux récepteurs œstrogènes et est par conséquent utilisé dans le cas du cancer du sein⁹⁶. La quatrième molécule du classement, le fosinopril, est utilisé dans le traitement de l'hypertension et de certains types d'insuffisance cardiaque. Il inhibe l'enzyme de conversion de l'angiotensine⁹⁷. Finalement, le demecarium est donné pour le traitement de glaucome. Il inhibe l'anticholinestérase⁹⁸. Parmi ces cinq composés, un seul a déjà été décrit dans le cadre du cancer et aucun n'est décrit comme se liant dans un site de liaison du NAD⁺. Il est intéressant de noter qu'une poche secondaire formée pour le SBD, NDB et la partie modélisée de l'enzyme est présente. Celle-ci présente six acides aminés formant préférentiellement des interactions avec les composés se liant dans cette poche : P98, P207, L379, K384, V388 et V391 (Figure 6-5).

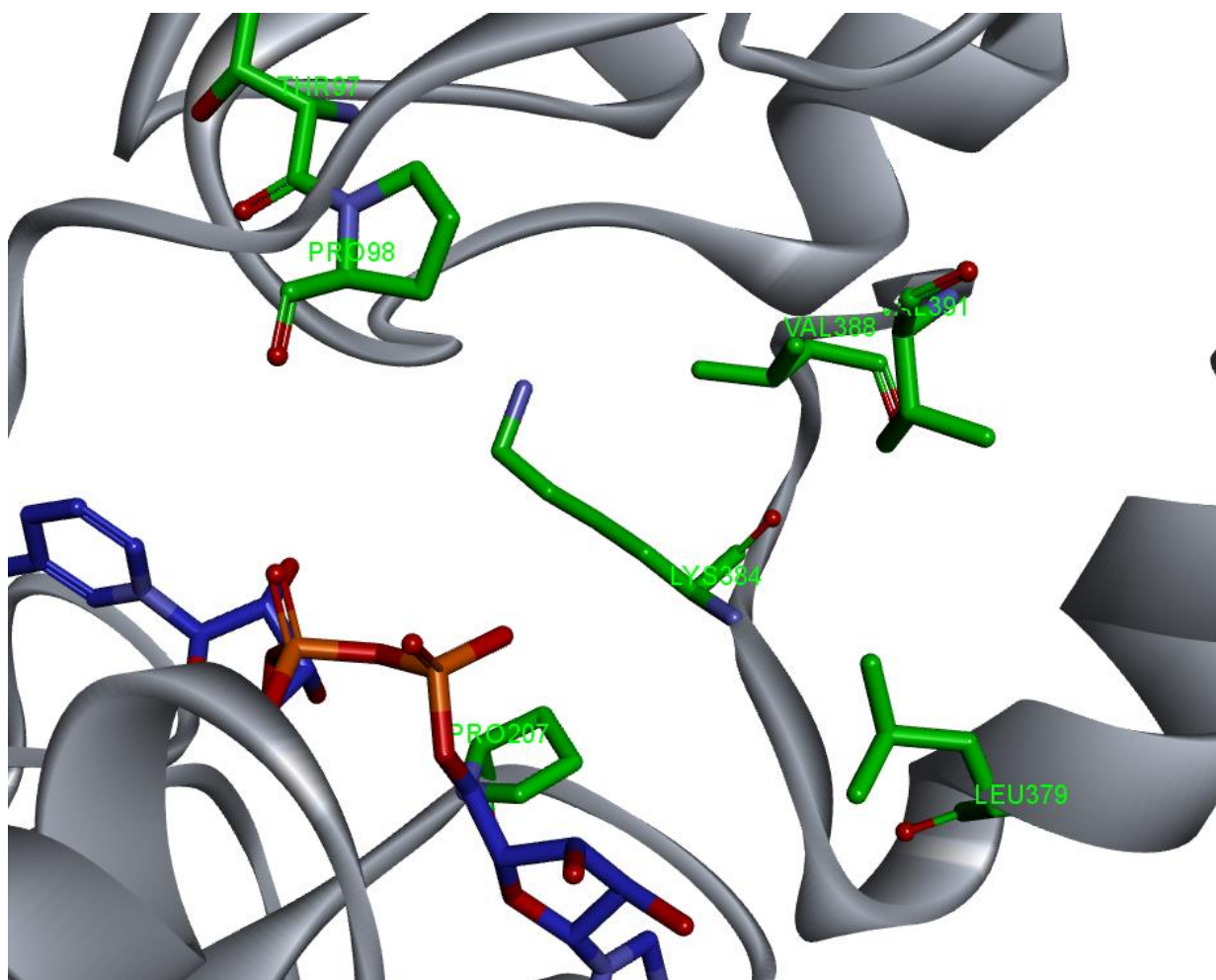


Figure 6-5: Poche secondaire présente dans la poche de liaison au cofacteur. Les acides aminés principaux la formant sont représentés en vert, le NAD⁺ est représenté en bleu et le squelette de l'enzyme est représenté en gris.

Les composés présentant les meilleurs scores ont privilégié la liaison au sein de la poche secondaire. Ils ont toutefois conservé pour la plupart un point d'ancrage dans la poche de liaison au NAD⁺ à la position de la liaison des deux groupements phosphates. Cette partie de la poche caractérisée par les acides aminés R154, I155 et E399 est aussi en interaction avec la majorité des composés obtenus avec l'ensemble « druglike ». Cet ensemble, présentant un plus grand nombre de composés, a permis d'obtenir des composés se liant au sein de la poche du NAD⁺. Nous y observons que les acides aminés D174, H205, C233 et R235 répartis dans la poche semblent aussi être des points d'interactions privilégiés en plus du site de liaison des deux groupements phosphates. Finalement, nous observons pour les composés présentant le meilleur score une liaison dans la poche secondaire précédemment observée. La présence d'une telle poche, bien

qu'elle demande à être confirmée pourrait être intéressante d'un point de vue thérapeutique. En effet, celle-ci permettrait d'utiliser des composés empêchant la liaison du cofacteur tout en présentant une sélectivité pour PHGDH par rapport aux autres enzymes utilisant aussi le NAD⁺ comme cofacteur. Remarquons également que les molécules **11** et **12** se positionnent dans la poche du NAD⁺ et du malate. Ce positionnement a l'avantage de présenter deux points d'ancrage. La molécule **12** interagit d'une part dans la partie de la poche liant les deux phosphates grâce à des interactions avec R154 et I155 et d'autre part dans la poche de liaison au malate grâce à R74, T77, N101, R235 et H282. Ce composé semble le plus intéressant au niveau de sa position dans la poche et des interactions qu'il forme. Sa conformation au sein du site actif en comparaison avec celle du NAD⁺ et du malate est représentée à la Figure 6-6. Nous y observons que si la liaison de ce composé s'avère correcte, il pourrait empêcher la liaison du NAD⁺ et du malate. De plus, la présence d'un groupement phényle se liant dans la poche secondaire pourrait permettre une meilleure sélectivité pour PHGDH.

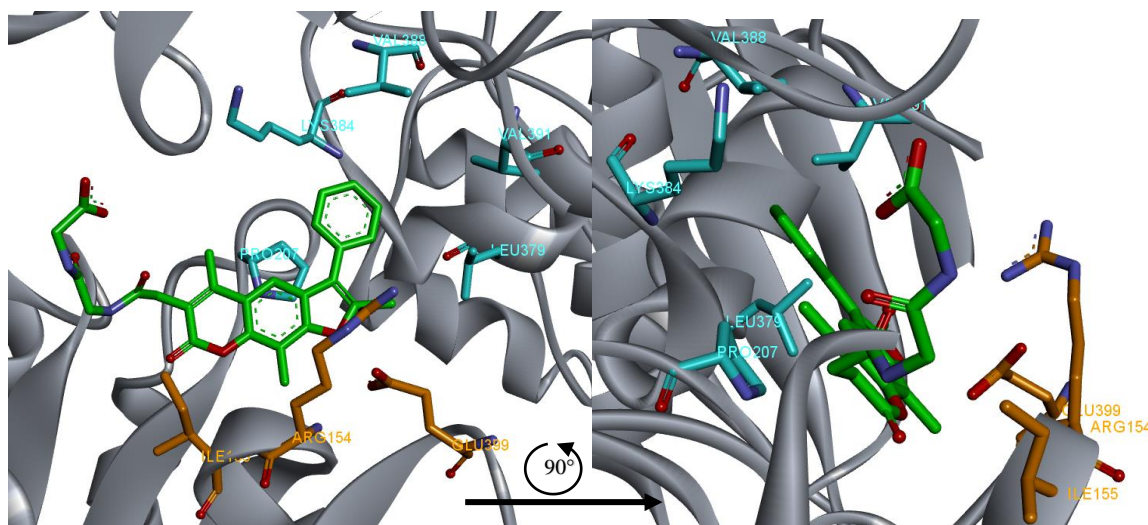


Figure 6-6: La conformation du 2-[[2-[[2-(7-keto-2,5,9-trimethyl-3-phenyl-furo[3,2-g]chromen-6-yl)acetyl]amino]acetyl]amino]acetate simulée par docking (en vert) au sein de la poche du NAD⁺ (en gris). Les trois acides aminés en interaction avec le plus grand nombre de molécules sont représentés en orange et les acides aminés composant la poche secondaire sont représentés en cyan.

6.3 Poche de liaison au substrat

La même procédure que pour la poche de liaison au NAD⁺ a été utilisée pour celle du 3-phosphoglycérate. Les résultats pour chaque ensemble sont présentés séparément et ensuite analysés dans la discussion.

6.3.1 Résultats obtenus pour l'ensemble « drug »

Les composés présentant le meilleur score lors de la simulation de leur conformation au sein du site actif sont présentés au Tableau 6-6. La conformation qui leur est associée est quant à elle représentée à la Figure 3-1Figure 6-7.

Tableau 6-6: Les cinq composés de l'ensemble "drug" dont la conformation générée dans la poche de liaison au substrat présente le meilleur score et le score associé.

N°	MOLÉCULE	SCORE
21	Acide folique	84
22	Alimta	80
23	D-(-)-Acide tartrique	78
24	L-(+)- Acide tartrique	78
25	Raltitrexed	78

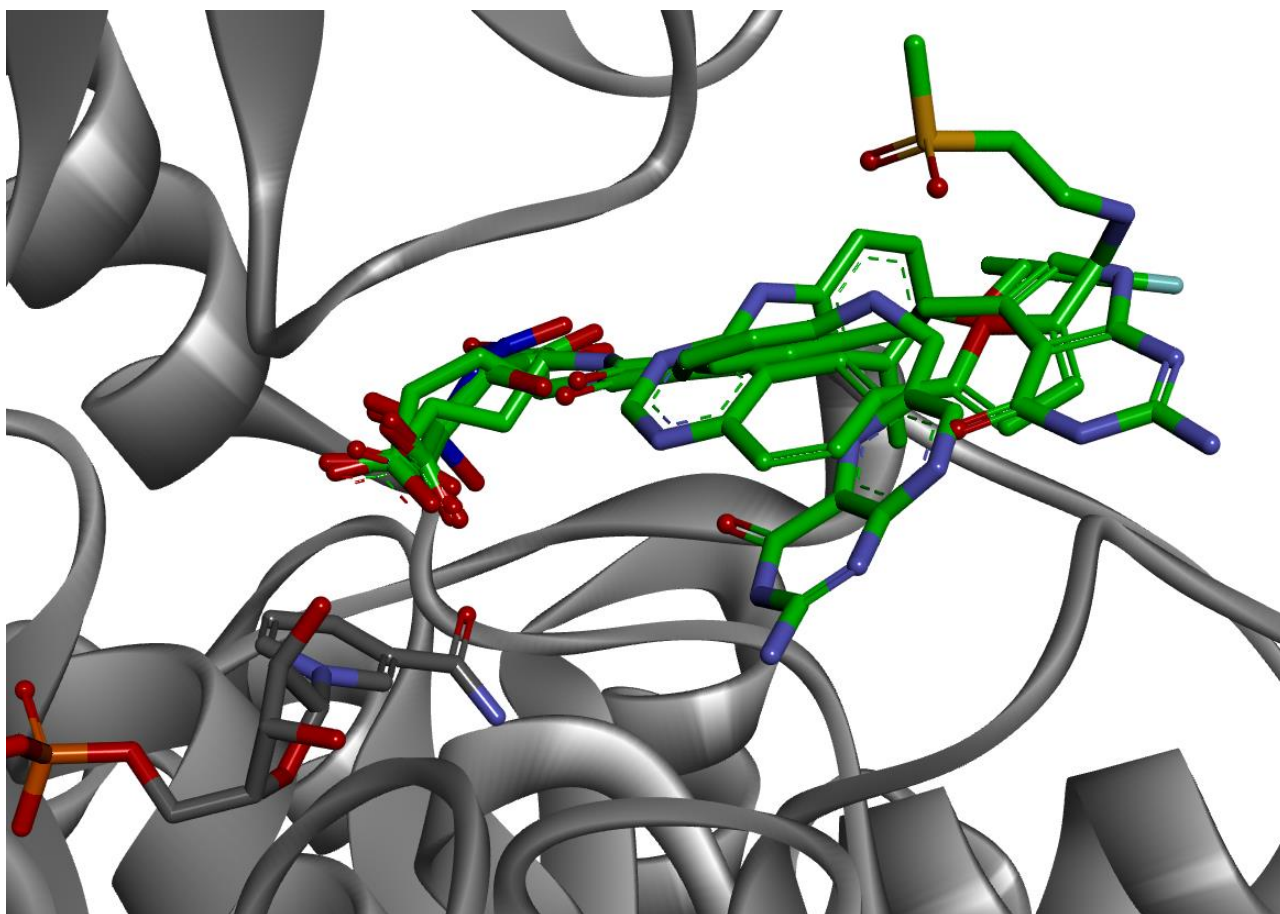


Figure 6-7: La conformation des cinq composés (en vert) obtenant le meilleur score par docking au sein de la poche du 3-phosphoglycérate (en gris) est comparée à la conformation du malate (en bleu).

Nous remarquons tout d'abord, dans le Tableau 6-6, la présence de l'acide tartrique dans les composés présentant le meilleur score. Ce composé a déjà été co-cristallisé dans le site actif de la forme tronquée de l'enzyme (5N6C)⁵⁵. Sur la Figure 6-7, nous observons que tous les composés présentent un acide carboxylique en bout de chaîne. Celui-ci se positionne au même endroit que celui du malate, en direction de la poche du NAD⁺. Pour toutes les molécules, cette fonction est stabilisée par pont hydrogène avec R74, T77, R235. Notons qu'en plus de ces trois acides aminés, R53, R134b et H282 sont eux aussi en interactions avec toutes les molécules. Un récapitulatif des interactions observées est repris au Tableau 6-10.

Le score de chaque composé a ensuite été normalisé en fonction de son nombre d'atomes lourds afin de générer un nouveau classement. Celui-ci est décrit au Tableau 6-7. Les conformations de chacune de ces molécules au sein du site actif sont représentées à la Figure 6-8.

Tableau 6-7: Les cinq composés de l'ensemble "drug" dont la conformation générée dans la poche de liaison au substrat présente le meilleur score normalisé et le score associé.

N°	MOLÉCULE	SCORE NORMALISÉ
26	Glycine	8,7
27	Oxybate de sodium	8,0
28	Acide 5-aminolevulinic hydrochloride	7,7
29	L-cystéine	7,3
30	Acide fluoropyruvique	7,3

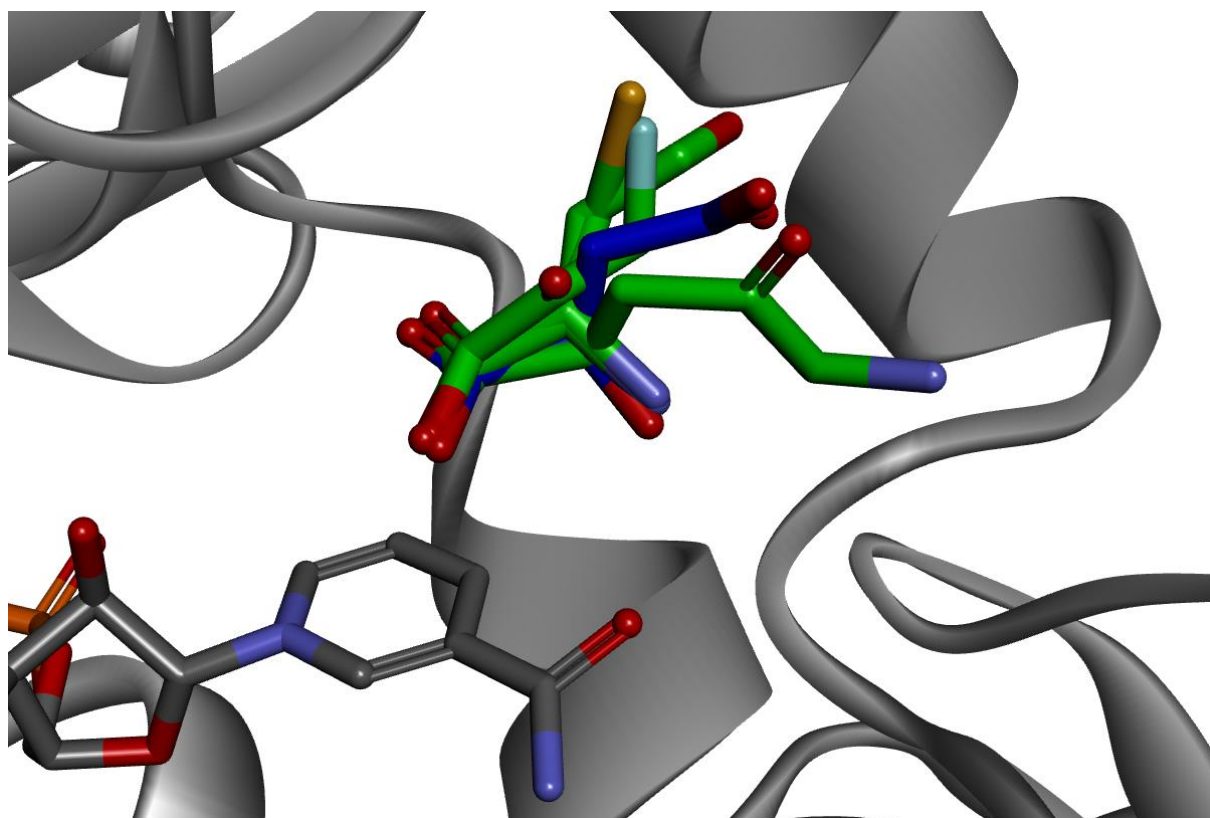


Figure 6-8: La conformation des cinq composés (en vert) obtenant le meilleur score normalisé par docking au sein de la poche du 3-phosphoglycérate (en gris) est comparée à la conformation du malate (en bleu).

Nous observons sur la Figure 6-8 que tous les composés présentent un mode de liaison similaire. En effet, nous retrouvons une interaction privilégiée entre une fonction acide carboxylique et les acides aminés T77, N101 et R235. On peut également noter que la L-cystéine et la glycine forment une interaction par pont hydrogène avec le NAD⁺.

Enfin, il est intéressant de souligner le fait que des composés (glycine, oxybate de sodium, L-alanine) qui avaient été observés dans la section 6.2.1 comme se liant dans le site de liaison du malate sont à nouveau observés. Le récapitulatif des acides aminés en interaction avec chaque composé est repris au Tableau 6-10.

6.3.2 Résultats obtenus pour l'ensemble « druglike »

Les 98179 molécules du second ensemble ont finalement été placées au sein du site de liaison du substrat. Les cinq molécules dont la conformation présente le meilleur score sont présentées au Tableau 6-8 et les conformations associées au score représentées à la Figure 6-9.

Tableau 6-8: Les 5 composés de l'ensemble "druglike" dont la conformation simulée dans la poche de liaison au substrat présente le meilleur score et le score associé.

N°	MOLÉCULE	SCORE
31	4-[2-(3-phenylpropanoylamino)benzoyl]aminobenzoic	77
32	4-[[[2-(2-keto-4,8-dimethyl-chromen-7-yl)oxyacetyl]amino]methyl]cyclohexanecarboxylate	72
33	(1S)-2-homoveratryl-1-(4-hydroxyphenyl)-7-methyl-1H-chromeno[2,3-c]pyrrole-3,9-quinone	72
34	2-[(4R)-2-[(6-ethoxy-4-methyl-quinazolin-2-yl)amino]-5-oxo-1,4-dihydroimidazol-4-yl]acetic	71
35	4-[[[2-[(4-keto-6-methyl-2,3-dihydro-1H-cyclopenta[c]chromen-7-yl)oxy]acetyl]amino]butyrate	71

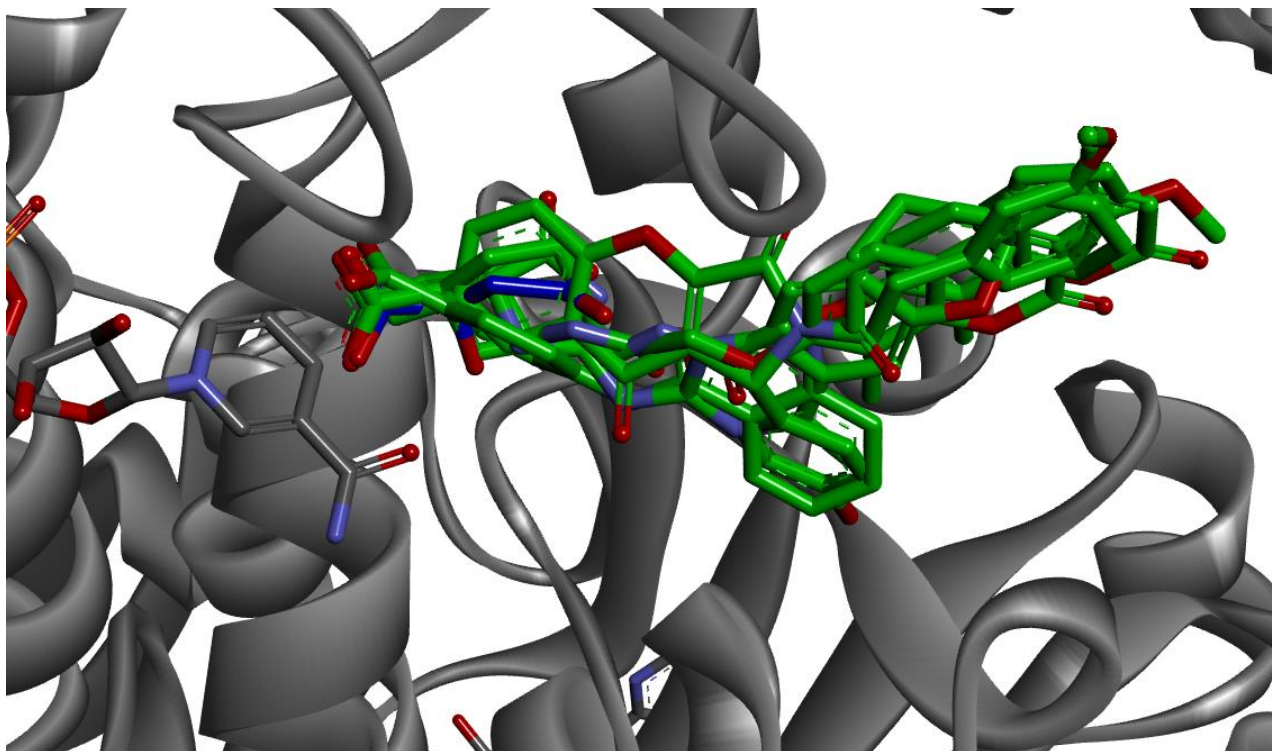


Figure 6-9: La conformation des cinq composés (en vert) obtenant le meilleur score par docking au sein de la poche du 3-phosphoglycérate (en gris) est comparée à la conformation du malate (en bleu).

Nous pouvons observer sur la Figure 6-9 que tous les composés privilégient une interaction au sein du site actif avec une fonction carboxylate tout comme le malate. Ce point d'ancrage privilégié est favorisé grâce à R74, T77, R235 et H282. Seule la molécule **33** ne présente pas cette fonction carboxylate mais une fonction chromeno capable de former des interactions entre les orbitales π de ses deux cycles et les acides aminés R53, S54, R74, H282. Tous les composés forment ensuite des interactions en surface de la poche catalytique. Ces interactions sont reprises dans le tableau récapitulatif présenté au Tableau 6-10.

Pour terminer, chaque composé a vu son score normalisé en fonction du nombre d'atome lourd de la molécule. Un nouveau classement a été mis en place et est présenté au Tableau 6-9 avec la conformation associée à chaque molécule à la Figure 6-10.

Tableau 6-9: Les cinq composés de l'ensemble "druglike" dont la conformation générée dans la poche de liaison au substrat présente le meilleur score normalisé et le score associé.

N°	MOLÉCULE	SCORE NORMALISÉ
36	Acide homogentisique	4,6
37	Gabapentin	4,5
38	4-(2-hydroxyethyl)benzene-1,3-diol	4,5
39	3-(carboxymethyl)-5-oxo-tetrahydrofuran-2-carboxylique	4,4
40	Danshensu de sodium	4,4

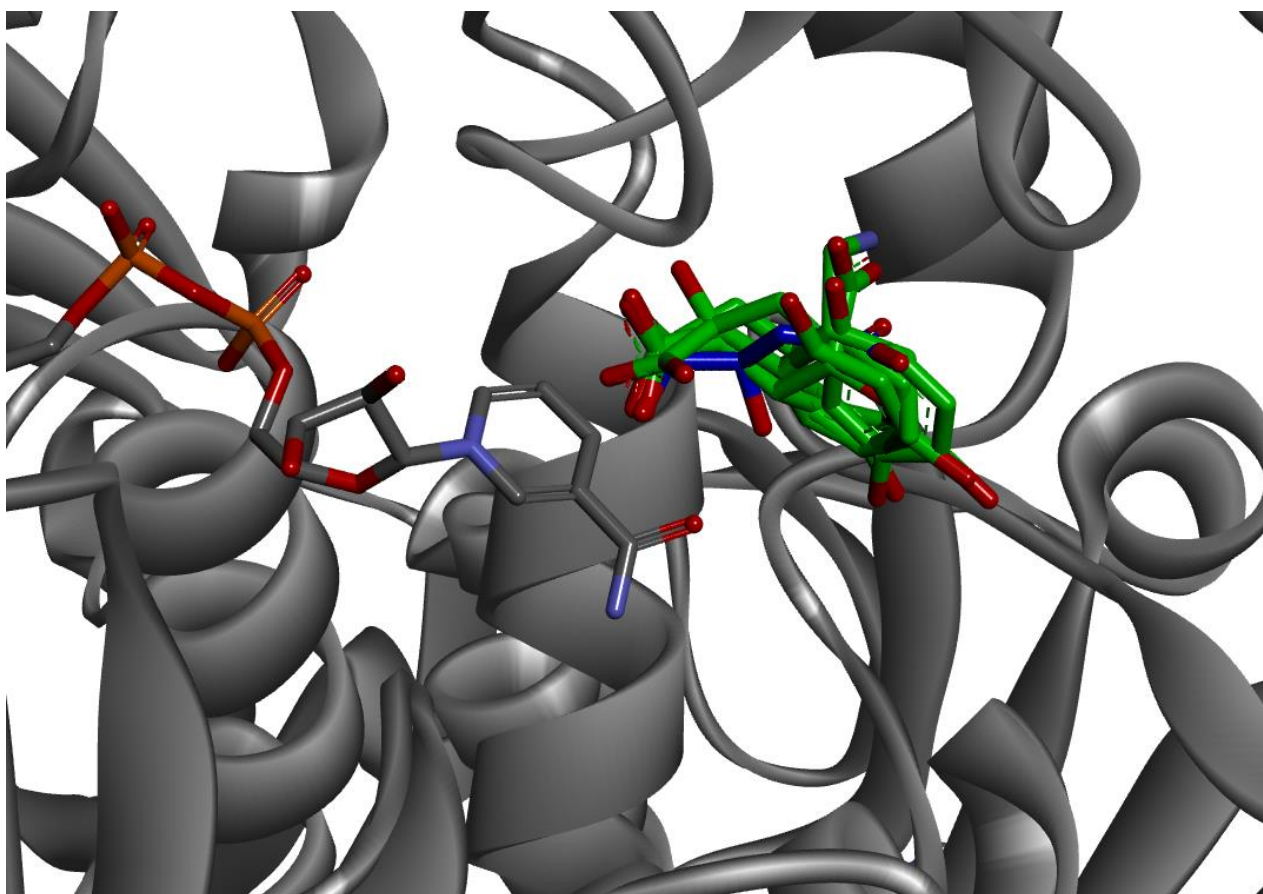


Figure 6-10: La conformation des 5 composés (en vert) obtenant le meilleur score normalisé par docking au sein de la poche du 3-phosphoglycérate (en gris) est comparée à la conformation du malate (en bleu).

Comme nous pouvons le voir sur la Figure 6-10, tous les composés présentent une fonction carboxylate (ou alcool dans le cas de la molécule **38**) permettant, comme indiqué précédemment, d'interagir avec le point d'ancrage formé par R53, S54, R74 et

H282. Nous remarquons également que les interactions formées par ces molécules sont très similaires à celles formées par le malate. Ces interactions sont représentées au Tableau 6-10 et le mode de liaison de chaque molécule est détaillé à l'Annexe C.

6.3.3 Discussion

Le Tableau 6-10 montre tous les acides aminés formant des interactions par pont hydrogène (vert) ou π -cation (orange) avec les molécules se liant dans la poche de liaison du malate. Remarquons que parmi les composés présentant le meilleur score de l'ensemble « drug », l'alimta est un analogue de l'acide folique et présente un mode de liaison similaire. De plus ce composé est déjà donné dans le cadre du cancer. Il est capable d'inhiber la thymidylate synthétase (TS), la dihydrofolate réductase (DHFR) et la glycinamide ribonucléotide formyltransférase⁹⁹. Le lapatinib est lui aussi un composé anti-cancéreux donné en cas de cancer du sein et de tumeurs solides. C'est un inhibiteur du récepteur de l'EGF et de la tyrosine kinase HER2¹⁰⁰.

Tableau 6-10: Récapitulatif des acides aminés formant des interactions par pont hydrogène (en vert) ou π -cation (en orange) avec les composés des ensembles « drug » et « druglike » dont la conformation a été générée dans la poche de liaison au malate.

MOLECULE	R53	S54	K57	R74	A75	T77	D80	N101	R134b	R235	G236	T262	E264	H282	A285	S286	Q291
21	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
22	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
23	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
24	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
25	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
26																	
27	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■
28	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
29	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
30	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
31																	
32	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
33	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
34	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
35	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
36	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
37	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
38	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
39	■	■	■	■	■	■	■	■	■	■	■	■	■	■			
40	■	■	■	■	■	■	■	■	■	■	■	■	■	■			

Nous pouvons classifier les acides aminés du Tableau 6-10 en 3 catégories. La première est constituée des acides aminés essentiels à la liaison d'un ligand et reprend tous les acides aminés se liant à 16 composés ou plus : R53, R74, T77 et R235. La seconde comprend les acides aminés se liant à un total de 10 à 15 composés : S54, N101 et R134b. Le reste des acides aminés se place dans la 3^{ème} catégorie, qui regroupe les acides aminés pouvant apporter une interaction supplémentaire mais qui ne sont pas primordiaux à la liaison d'un ligand. La Figure 6-11 montre la position de ces acides aminés dans le site actif. Nous pouvons voir que, comme mis en évidence dans la description des interactions des ligands, le carboxylate interagissant dans le fond de la poche est d'une grande importance pour toute molécule se plaçant dans cette poche. La stabilisation observée pour la fonction phosphate (dans le cas du 3-PG) et carboxylate (pour le malate) montre une plus grande flexibilité des fonctions pouvant se lier à cette position. Nous y retrouvons majoritairement des acides aminés de catégorie 2 et nous avons observé lors de notre criblage virtuel une plus grande variété de fonctions capables de se lier dans cette partie de la poche. Finalement, bien que des interactions plus en surface de la poche soient possibles, peu de ligands parmi ceux que nous avons étudiés semblent en former.

La Figure 6-11 illustre la conformation du 3-phosphoglycérate au sein du site actif et met en évidence que les catégories 1 (en vert) et 2 (en jaune) d'acides aminés correspondent au mode de liaison du substrat. Sur les 40 molécules étudiées, seules deux molécules présentent des interactions avec tous ces acides aminés : les deux isomères du tartrate. Cette molécule étant une des deux co-cristallisées au sein du site actif, cela confirme que ces interactions y sont importantes pour la liaison des isomères.

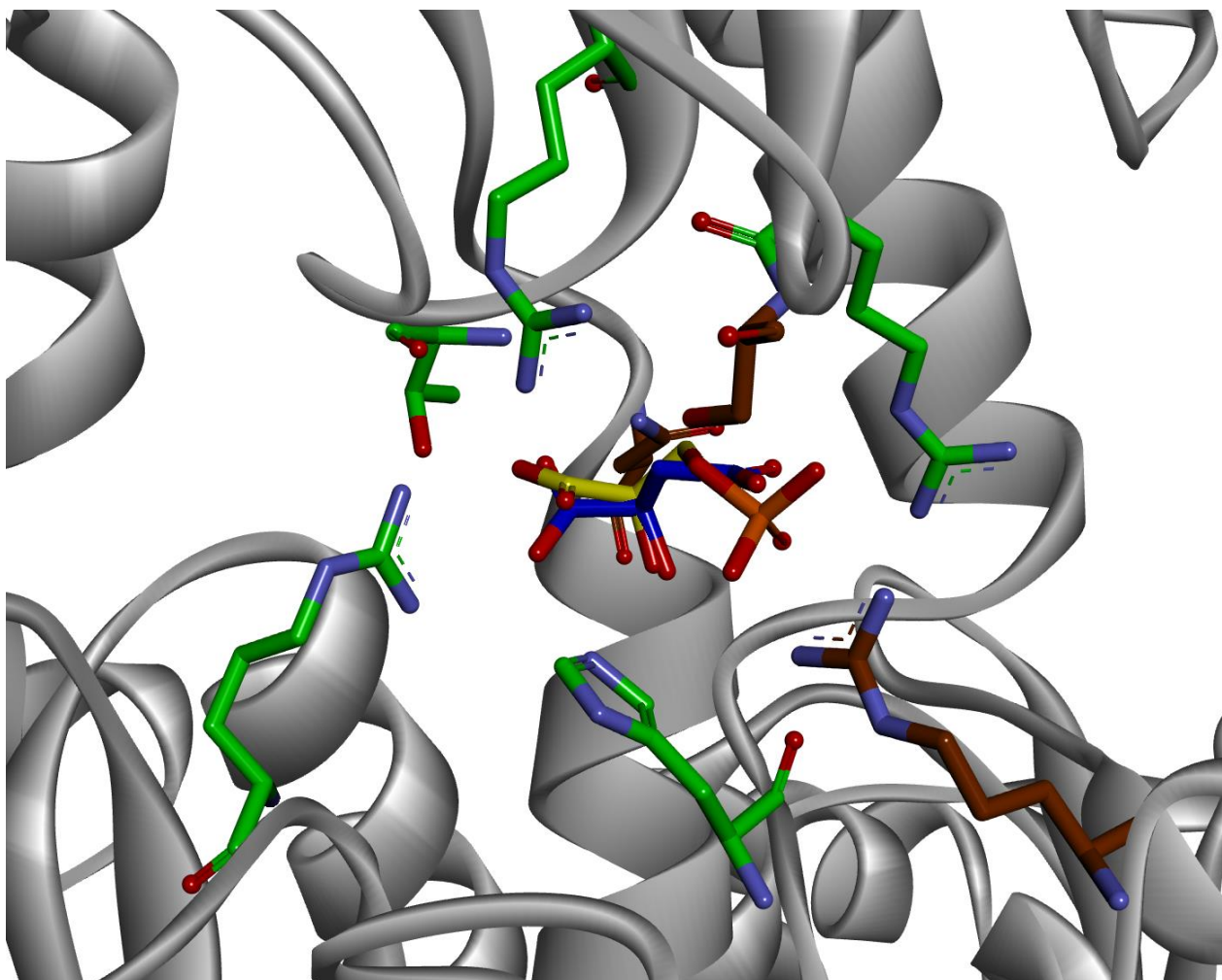


Figure 6-11: Poche de liaison de PHGDH en présence du malate (2G76, en bleu) et du 3-phosphoglycérate (docking, en jaune). Les acides aminés de catégorie 1 (R53, R74, T77 et R235, en vert) et de catégorie 2 (S54, N101 et R134b, en brun) sont représentés sous forme de bâtonnets.

6.4 L-cystéine et fluoropyruvate : deux ligands de la poche de liaison au substrat

Parmi les 40 composés présentés à la section précédente, deux ont été sélectionnés afin d'être co-cristallisés avec la forme tronquée de l'enzyme : la L-cystéine et le fluoropyruvate. Le détail de production et purification de la protéine ainsi que l'obtention des cristaux sont présentés au chapitre 1.

6.4.1 Cristallisation du fluoropyruvate au sein du site actif

Le cristal de sPHGDH a été obtenu par goutte pendante en ratio 1 : 1 entre l'enzyme (15 mg/mL) et la solution d'un réservoir contenant 0,1M Bistris (pH = 5,5) ; 0,2M Li₂SO₄ ; 24% PEG 3350. La solution a été incubée 2 heures en présence de fluoropyruvate à température ambiante. Les données ont été collectées au synchrotron de SOLEIL (Gif-sur-Yvette, France, lignes PROXIMA 1 et 2). Les statistiques de l'affinement et des données collectées sont présentées au Tableau 6-11.

Tableau 6-11: Données collectées et statistiques de l'affinement (déterminée avec Molprobit), les valeurs entre parenthèses sont celles à la plus haute résolution.

SPHGDH FLUOROPYRUVATE	
DONNÉES DU CRISTAL	
Groupe d'espace	P1
Dimensions	
a , b , c (Å)	43,22 ; 59,70 ; 64,76
α , β , γ (°)	88,81 ; 90.04 ; 80.13
Résolution (Å)	44.00-2.20 (2,28 - 2,20)
Nombre de réflexions uniques	20521 (1023)
Complétude (%)	63.14% (31.33%)
I/ σ I	3,3 (0,4)
Redondance	5,4 (2,5)
STATISTIQUES DE L'AFFINEMENT	
Ligand	NAD+
Rwork / Rfree (%)	0.2068/0.2385
Nombre d'atomes	4644
Macromolécules	4463
N° molécules d'eau	30
RMS (liaison) (Å)	0,0138
RMS (angle) (°)	1,75
Ramachandran favorable (%)	94,32

La carte de densité électronique centrée sur le site de liaison au substrat est présentée à la Figure 6-12. Nous y voyons la présence de densité électronique correspondant au fluoropyruvate. Cependant, comme nous pouvons le voir sur la droite de l'image le domaine de liaison au substrat, flexible, est mal défini. Ce manque d'informations au sein du site de liaison ne permet pas d'établir une carte de densité suffisamment définie pour le fluoropyruvate et par conséquent de confirmer son mode de liaison. Sur la Figure 6-13, les deux monomères composants le dimère ont été alignés et nous pouvons voir que la position du fluoropyruvate est différente, ce qui atteste de la difficulté de confirmer un mode de liaison pour ce ligand.

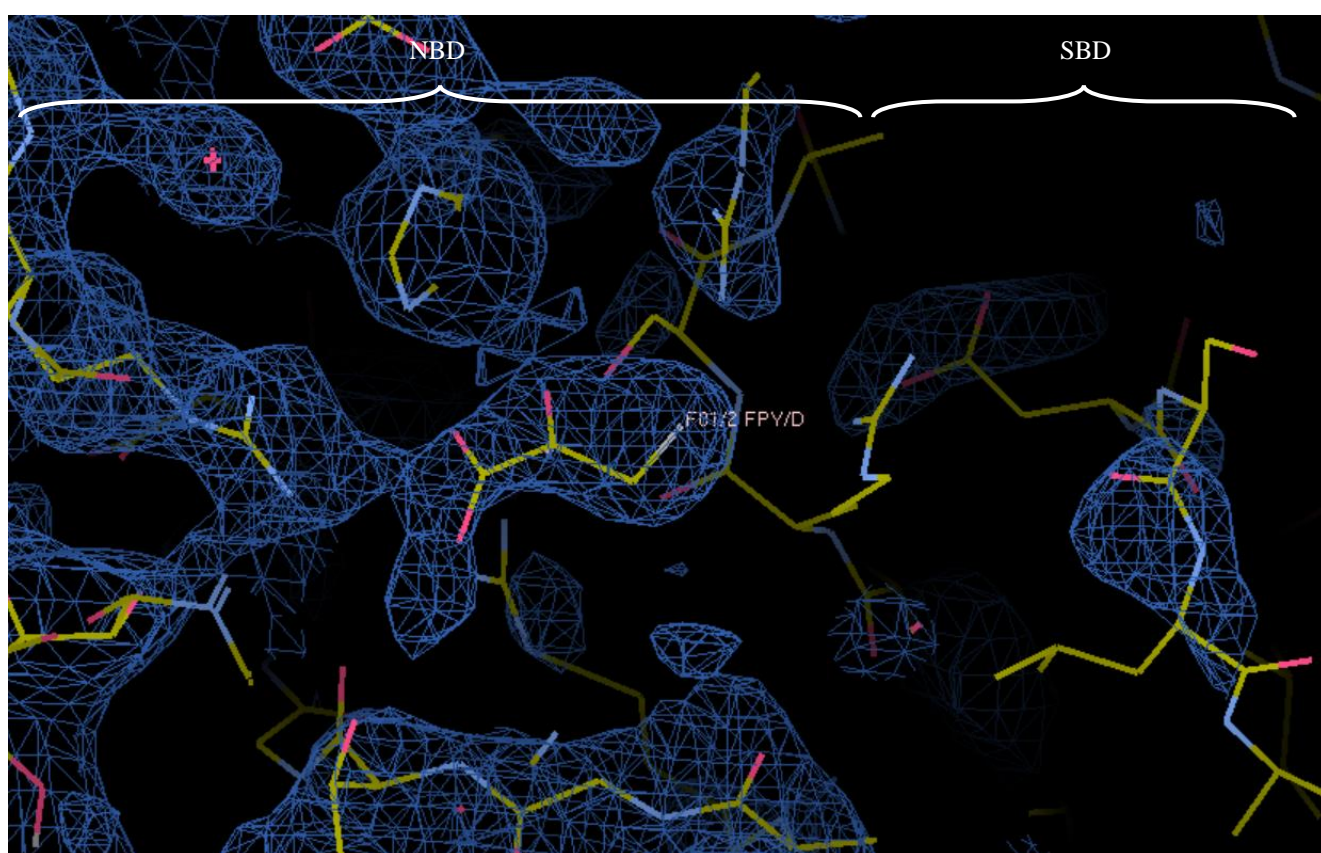


Figure 6-12: Carte de densité électronique ($2Fo-Fc$, $RMS\ 0,2e/\text{\AA}^3$, en bleu) après affinement représentant le site de liaison au substrat au centre de celui-ci, le ligand co-cristallisé: le fluoropyruvate.

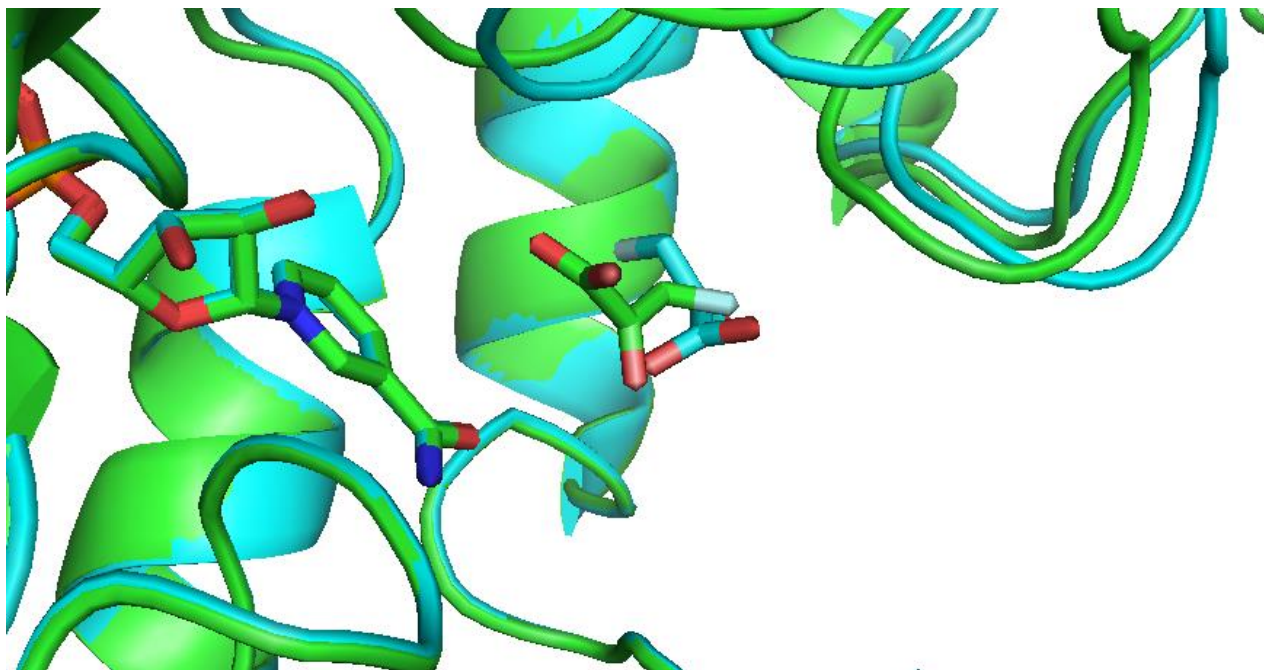


Figure 6-13: Alignement des deux monomères (en vert et en cyan) présent dans l'unité asymétrique lors de la co-cristallisation du fluoropyruvate. Les ligands sont présentés sous forme de bâtonnet.

La structure globale du dimère est quant à elle très semblable à celle de sPHGDH déjà publiée (2G76). Un RMSD de 0,55Å entre les carbones alphas est calculé après alignement des deux structures, et la présence du cofacteur NAD⁺ est aussi observée dans notre structure.

6.4.2 Cristallisation de la L-cystéine au sein du site actif

La cristallisation de la forme tronquée de PHGDH a été réalisée en goutte pendante avec un ratio 1 :2 entre l'enzyme (15 mg/mL) et la solution d'un réservoir contenant 0,1M Bistris (pH= 6,5) ; 0,2M Li₂SO₄ ; 22% PEG 3350. La solution a été incubée 2 heures en présence de L-cystéine. Les cristaux ont été générés et les données acquises au synchrotron de SOLEIL (Gif-sur-Yvette, France). Les statistiques de l'affinement et des données collectées sont présentées au Tableau 6-12.

Tableau 6-12: Données collectées et statistiques de l'affinement (déterminée avec Molprobit), les valeurs entre parenthèses sont celles à la plus haute résolution.

SPHGDH L-CYSTÉINE	
DONNÉES DU CRISTAL	
Groupe d'espace	P1
Dimensions	
a , b , c (Å)	43,22 , 59,70 , 64,76
α , β , γ (°)	88,81 , 90,04 , 80,13
Résolution (Å)	44,00-2,18 (2,27 - 2,2)
Nombre de réflexions uniques	20521 (1022)
Complétude (%)	63.14% (31.33%)
I/ σ I	4,0 (0,5)
Redondance	7,5 (0,9)
STATISTIQUES DE L'AFFINEMENT	
Ligand	NAD+
Rwork / Rfree (%)	0.1988/0.2294
Nombre d'atomes	4644
Macromolécules	4463
N° molécules d'eau	30
RMS (liaison) (Å)	0,0138
RMS (angle) (°)	1,75
Ramachandran favorable (%)	94,32%

De la même façon que pour la carte de densité électronique du fluoropyruvate, nous observons une faible densité électronique pour le domaine de liaison au substrat, possiblement dû à sa flexibilité. Comme pour le fluoropyruvate, la présence d'une densité au sein du site actif atteste de la présence d'un ligand et correspond à la densité d'une molécule de L-cystéine (Figure 6-14). Cependant, celle-ci n'est pas suffisamment

bien définie que pour préciser de façon exacte la position de la L-cystéine au sein du site actif.

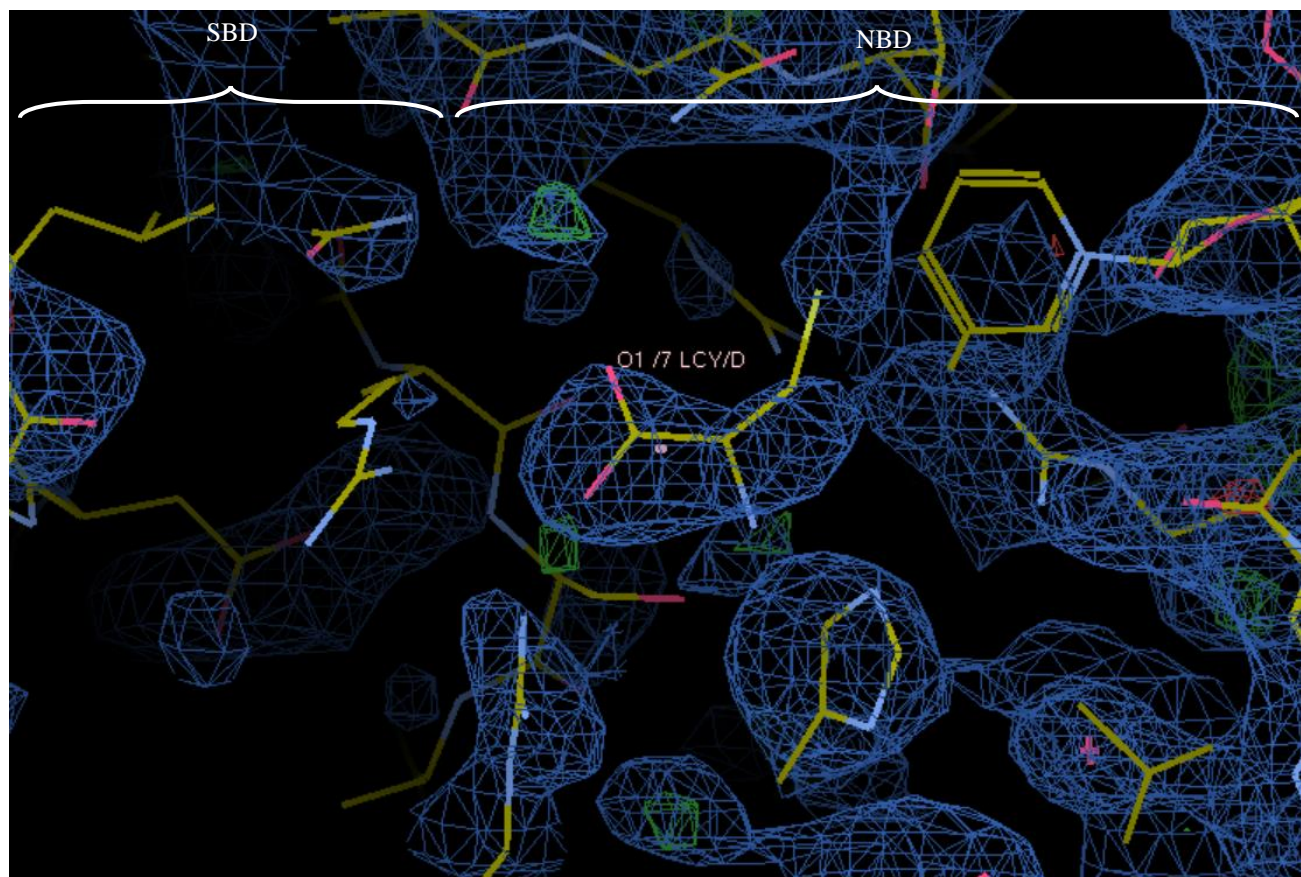


Figure 6-14: Carte de densité électronique ($2F_o - F_c$, $RMS\ 0,2e/\text{\AA}^3$, en bleu) après affinement représentant le site de liaison au substrat au centre de celui-ci, le ligand co-cristallisé : la L-cystéine.

Un alignement des deux monomères présents dans l'unité asymétrique avec un agrandissement centré sur la position de la L-cystéine est présenté à la Figure 6-15. Nous y observons à nouveau deux positions différentes attestant de la difficulté à déterminer le bon mode de liaison du substrat. Cette difficulté résulte de la faible définition de la densité électronique.

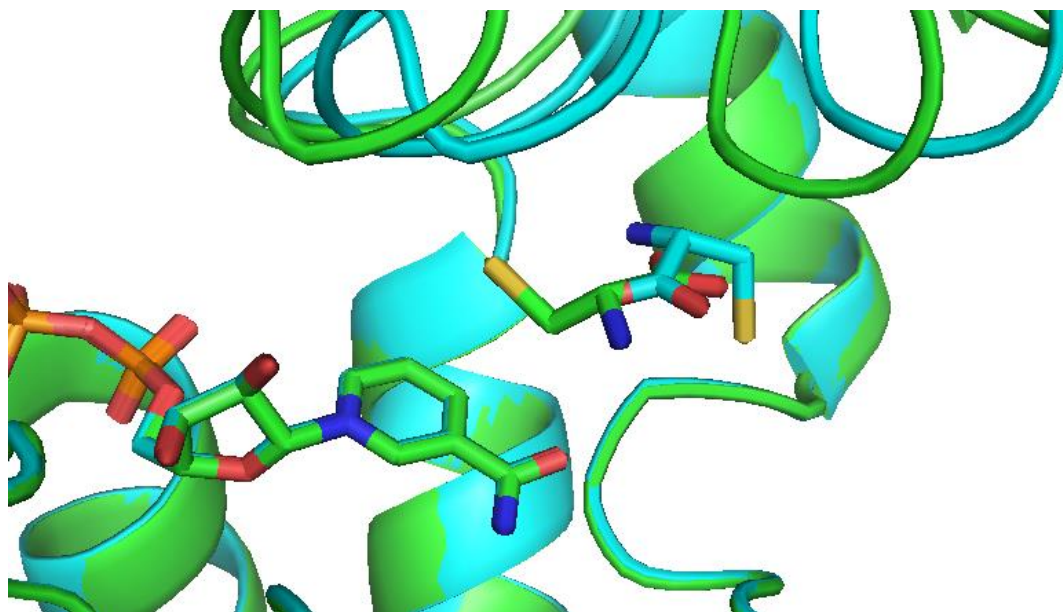


Figure 6-15: Alignement des deux monomères (en vert et en cyan) présent dans l'unité asymétrique lors de la co-cristallisation de la L-cystéine. Les ligands sont présentés sous forme de bâtonnets.

A nouveau, le NAD⁺ est présent et la structure globale de l'enzyme est très proche de celle déjà existante (2G76) avec un RMSD de 0,19 Å. Bien que les structures cristallographiques ne permettent pas de positionner avec précision le fluoropyruvate et la L-cystéine, le travail confirme que les deux ligands issus du docking peuvent se lier à la protéine.

6.5 Conclusion

Au terme de ce criblage, nous avons pu, grâce à la génération de conformations de plus de 100.000 molécules au sein de la poche de liaison au cofacteur et au substrat, mettre en évidence les sites d'ancrage essentiels à la liaison d'un ligand au sein de ces poches. La poche de liaison du NAD⁺ est une poche relativement volumineuse ne permettant pas la stabilisation de petites molécules. La partie de la poche qui permet la liaison des deux groupements phosphates du NAD⁺, formée par R154, I155 et E399, est essentielle pour la liaison de ligands au sein du site de liaison au NAD⁺. Les composés sont stabilisés par des interactions par ponts hydrogène, π -cation et π -alkyl, ce qui permet de lier différents types de *scaffold* (carboxylate, alcool, dihydrofurane, ...). Nous avons

aussi mis en évidence la présence d'une poche secondaire (P98, P207, L379, K384, V388 et V391) qui, si elle correctement modélisée, permettrait la stabilisation de fonctions alkyles et assurerait la spécificité des inhibiteurs vis-à-vis des autres enzymes NAD dépendantes (Figure 6-16). La molécule **12** (2-[[2-[[2-(7-keto-2,5,9-trimethyl-3-phenyl-furo[3,2-g]chromen-6-yl)acetyl]amino]acetyl]amino]acetate) semble la plus intéressante de par sa capacité à lier le site de liaison du cofacteur et du substrat tout en interagissant avec la poche secondaire par un groupement phényle. Il serait intéressant de tester le pouvoir inhibiteur de cette molécule avec et sans le groupement phényle. Si cette molécule est capable d'occuper la position du NAD⁺, nous pourrions dès lors savoir si la présence du phényle permet une meilleure inhibition (attestant de la présence de la poche secondaire) ou bien si celui-ci a un effet délétère sur l'inhibition de l'enzyme.

La génération de composés au sein de la poche de liaison au substrat a montré qu'il est nécessaire pour un bon ligand de présenter de nombreuses fonctions capables de former des ponts hydrogène avec les deux points d'ancrage mis en évidence. Les acides aminés que nous avons classés en catégorie 1 (R53, R74, T77 et R235) sont essentiels pour la liaison d'un ligand au sein du site actif (remarquons que la molécule **12** présente des interactions avec trois de ces quatre acides aminés) (Figure 6-16). Ces acides aminés présentent principalement des interactions par pont hydrogène mais peuvent aussi (principalement R54) stabiliser le ligand par interaction π -cation. Les acides aminés de catégorie 2 (S54, N101 et R134b) assurent un second point d'ancrage au sein du site actif. Nous observons un mode de liaison similaire pour tous les composés. La présence d'une fonction acide carboxylique pour interagir avec les acides aminés de catégorie 1 est essentielle (à l'exception de la molécule **38**, tous les autres ligands présentent un carboxylate à cette position). La présence du tartrate dans les molécules se liant au sein du site actif montre que la méthode utilisée permet d'identifier des composés déjà connus comme étant des ligands de PHGDH⁵⁵. La Figure 6-16 montre que qu'il est envisageable de concevoir des composés capables d'interagir aussi bien avec la poche de liaison du substrat et du cofacteur (en plus de la poche secondaire qui permettrait une meilleure sélectivité) comme observé pour le composé **12**.

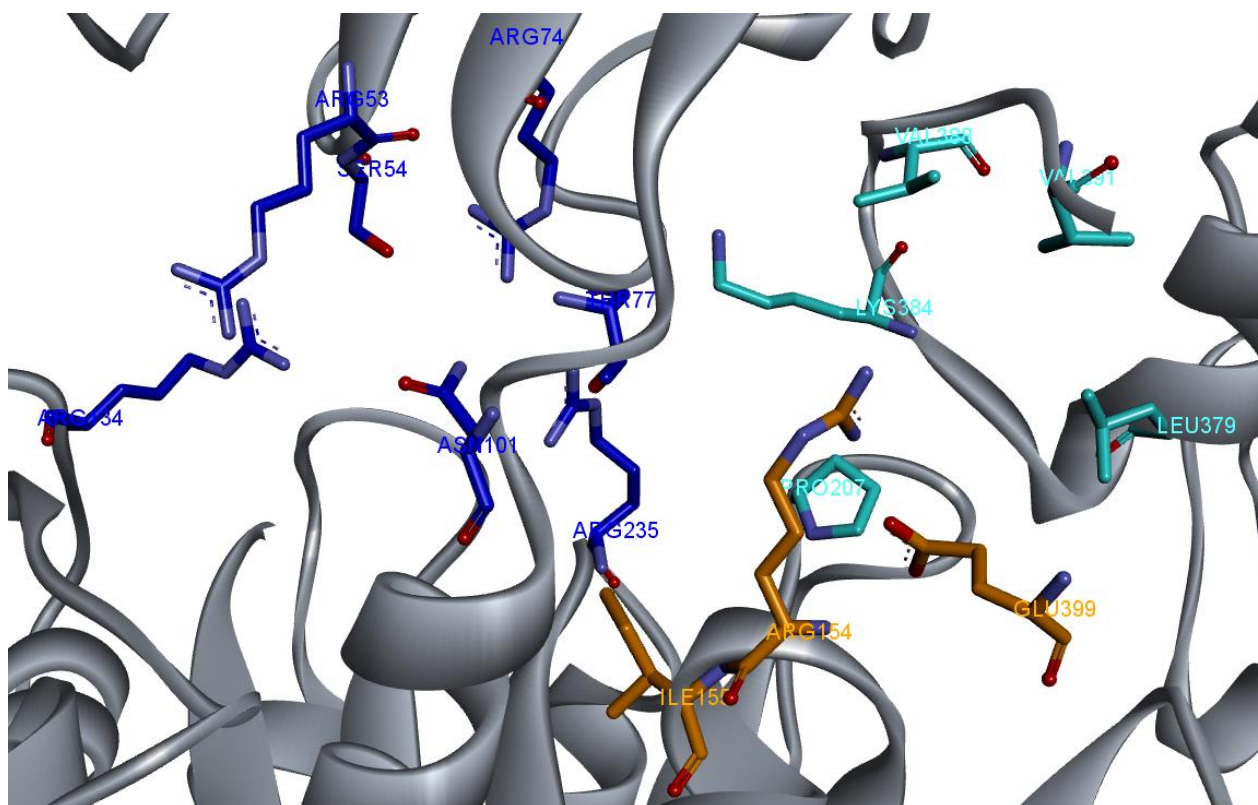


Figure 6-16: Site actif de la structure tétramérique de PHGDH modélisé. Les acides aminés mis en évidence lors du screening virtuel sont représentés sous forme de bâtonnets. Les acides aminés de la poche du substrat sont en bleus, ceux de la poche du cofacteur en orange et ceux de la poche secondaire en cyan.

Les informations apportées par les molécules identifiées ont permis d'identifier les sites d'interactions et les *scaffolds* se liant à ceux-ci. Il serait dès lors intéressant de tester certaines de ces molécules (tests d'activité et co-cristallisation) afin de confirmer les modes de liaison observés. Il serait aussi possible de concevoir de nouveaux ligands se liant au point d'ancrage du site de liaison du NAD⁺ et à ceux du malate à l'aide des *scaffolds* proposés.

Parmi les composés identifiés, trois sont déjà donnés dans le cadre de thérapie anti-cancéreuses : le fuvelstrant, le lapatinib et l'alimta. Si ces composés sont capables d'inhiber PHGDH, il serait possible d'élargir leur utilisation. Il serait intéressant d'étudier plus en détail les cancers pour lesquels ils sont utilisés et de voir si une meilleure inhibition est observée pour ceux ayant une haute expression en PHGDH.

La co-cristallisation a été réalisée avec deux composés se liant dans la poche du substrat : le fluoropyruvate et la L-cystéine. Bien que la qualité des données collectées

n'ait pas permis de préciser le mode de liaison de ces molécules, nous pouvons voir que celles-ci se positionnent effectivement au sein du site actif, ce qui confirme les résultats obtenus lors du *docking*.

7. Conclusion

7 Conclusion

7.1 Discussion

Depuis l'identification de l'implication de PHGDH dans des mécanismes cancéreux par Possemato⁵³ et Locasale⁵², de nombreuses recherches sur cette enzyme ont été réalisées, permettant l'identification de nouveaux ligands et inhibiteurs^{51,55,60,67-69}. Malheureusement, peu de ces inhibiteurs/ligands ont un mode de liaison connu avec l'enzyme et aucun n'est compétitif du site actif. Après la rédaction de ce manuscrit, *Mullarky et al.* ont publié une série de composés se liant dans la poche du NAD⁺ et co-cristallisé ceux-ci (2 nouvelles structures) au sein du site de liaison du cofacteur avec l'enzyme tronquée⁵⁹. Cependant, les informations structurales et fonctionnelles ciblant le site actif de l'enzyme sont rares tandis que celles traitant de l'enzyme complète et sous forme tétramérique sont inexistantes.

L'objectif de ce travail étant de fournir des informations structurales sur la forme complète de l'enzyme et des informations structurales et fonctionnelles ciblant le site actif de l'enzyme, nous avons logiquement commencé par la production et la cristallisation de l'enzyme. Le chapitre 1 met en évidence la difficulté de cristalliser PHGDH sous sa forme complète, ce qui explique le manque d'information structurale pour cette enzyme. Malgré l'utilisation d'une protéine fusionnée à une *Maltose Binding Protein* et la réduction d'entropie en surface de la protéine, nous n'avons jamais pu obtenir de cristaux diffractant suffisamment bien que pour permettre d'obtenir des données exploitables. Les domaines ASB et ACT responsables de la tétramérisation de l'enzyme semblent empêcher sa cristallisation. Plusieurs solutions sont envisageables afin d'obtenir des informations structurales complémentaires. Les domaines ASB/ACT pourraient être exprimés et cristallisés seuls. D'autres techniques structurales telle que la résonance magnétique nucléaire sont également envisageables. Cependant, son utilisation est limitée à des enzymes de petites tailles¹⁰¹. La technique la plus prometteuse serait la cryo-EM qui ne nécessite pas la cristallisation de l'enzyme.

Afin de pouvoir étudier l'effet de la mutation R134A et le mode de liaison de la L-cystéine et du fluoropyruvate, nous avons utilisé la forme tronquée de l'enzyme, seule forme de l'enzyme ayant déjà permis l'obtention d'une structure cristallographique. Des données exploitables ont été obtenues et ont permis la résolution de la structure de l'enzyme. Nous avons pu observer dans les différentes structures la mutation R134A et la liaison de la L-cystéine et du fluoropyruvate au sein du site de liaison du substrat. Malheureusement, la carte de densité du domaine de liaison au substrat n'étant pas suffisamment définie à cause de la flexibilité de celui-ci, elle n'a pas permis de définir le mode de liaison exacte de ces deux molécules.

Afin d'obtenir des informations structurales sur l'enzyme complète, nous avons généré un modèle de celle-ci. Deux programmes de modélisation indépendants ont été utilisés (SWISS MODEL et i-TASSER) et ont permis d'obtenir deux structures similaires de l'enzyme. A l'aide du programme SWISS MODEL, nous avons pu générer la structure de PHGDH sous sa forme tétramérique (Figure 7-1). Ce modèle basé sur la forme bactérienne de PHGDH provenant de l'organisme *Mycobacterium tuberculosis* est la première structure de l'enzyme complète sous sa forme tétramérique. Ce modèle est supporté par des expérimentations de spectrométrie de masse couplées au *cross-linking*. Celles-ci ont permis de confirmer des zones d'interactions des différents monomères supportant la conformation globale du tétramère. Nous savons que PHGDH est active sous forme tétramérique et qu'une activité plus faible est observée pour la forme dimérique de l'enzyme. Sur base de ce modèle nous pouvons expliquer une perte d'affinité pour le cofacteur comme observé par notre collaborateur, le Dr. E. Mullarky, avec un $K_D(\text{NAD}^+) = 0,360 \text{ mM}$ dans le cas de l'enzyme complète et $K_D(\text{NAD}^+) = 65 \text{ mM}$ pour l'enzyme tronquée. Toutefois, il est difficile de dire si la perte drastique d'activité observée entre les deux formes de l'enzyme peut être attribuée uniquement à cette perte d'affinité, ou si d'autres mécanismes encore inconnus sont également impliqués. Une hypothèse serait que seulement deux des quatre sites actifs seraient actifs et catalyseraient la réaction tandis que les deux autres seraient inactifs et qu'un changement conformationnel permet d'inverser l'activité des différents sites. Ce type de mécanismes n'a toutefois jamais été décrit dans le cas des PHGDH de type I.

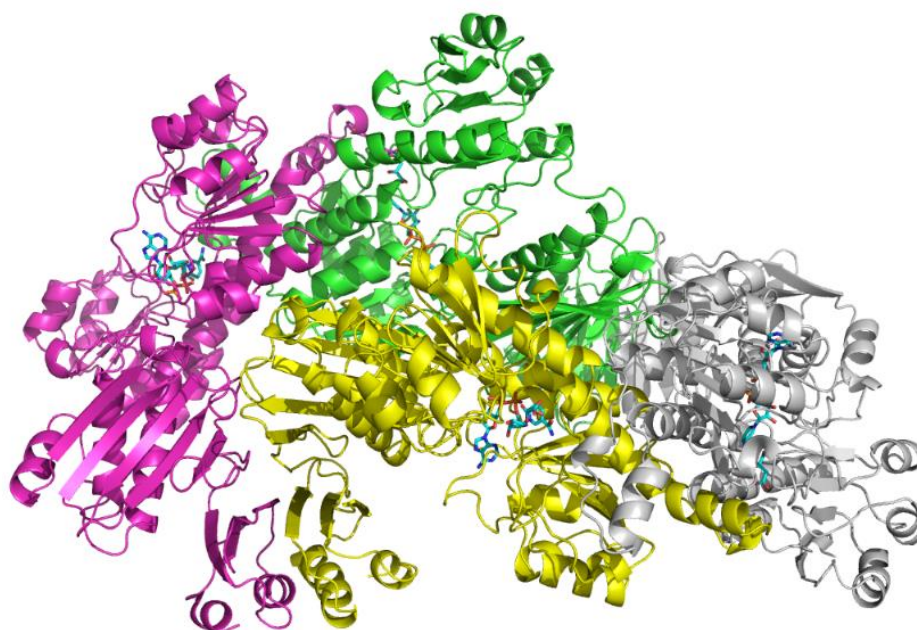


Figure 7-1: Modèle hybride de l'enzyme PHGDH complète sous sa forme tétramérique (chaque monomère est représenté dans une couleur différente).

Le modèle de l'enzyme généré (Figure 7-1) a permis de mettre en évidence le mode de liaison du 3-phosphoglycérate au sein du site actif. Nous y observons deux points d'ancrage. Le premier, se trouvant au niveau du carboxylate est stabilisé par R74, T77 et R235. Le second apparaît au niveau du groupement phosphate qui est en interaction avec S54, H282 et R134b. Finalement, le carbonyle qui est oxydé est orienté vers le NAD⁺ par interaction avec celui-ci et H282. Nous avons aussi investigué le rôle de R134b et avons mis en évidence que celui-ci est en effet impliqué dans la stabilisation du 3-PG au sein de la poche. Nous observons en effet un K_m de $0,107 \pm 0,032$ mM pour l'enzyme sauvage qui est plus de trois fois supérieur pour l'enzyme mutée ($0,337 \pm 0,098$ mM). Au-delà de cette observation, nous avons aussi remarqué une perte d'activité complète de l'enzyme lorsque la boucle contenant R134A est mutée. Celle-ci ne semble pas impliquée dans la formation du tétramère de l'enzyme et aucun des acides aminés mutés ne forme d'interaction avec un autre monomère de l'enzyme. Il serait intéressant de muter sélectivement un ou plusieurs acides aminés de cette boucle afin de comprendre l'impact de ces mutations sur l'activité et la structure de l'enzyme. Identifier le(s) acide(s) aminé(s) responsable(s) de cette perte d'activité permettrait d'envisager la conception d'inhibiteurs non sélectifs du site actif et entraînant une perte d'activité totale

de l'enzyme. Nous avons pu montrer qu'une légère perte d'activité de l'enzyme était observée en présence de peptides semblables à cette boucle, ce qui supporte l'intérêt de cette voie d'un point de vue thérapeutique. Les informations actuellement à notre disposition ne nous permettent pas de comprendre pourquoi la mutation de cette boucle entraîne une perte complète de l'activité de l'enzyme.

Un criblage virtuel sur plus de 100.000 composés dans les deux poches composant le site actif a également été réalisé. Ceci a permis d'identifier trois sites d'ancrage formant principalement des interactions par ponts hydrogène, π -cation ou π -alkyle. Ces points d'ancrage sont composés des acides aminés : R53, S54, R74, T77, N101, R134b, R154, I155, R235 et E399 (Figure 7-2). Interagir avec ces acides aminés garantit une bonne liaison au sein du site actif. Les trois sites d'ancrage correspondent d'ailleurs à des sites de liaison importants des ligands naturels : les deux groupements phosphate du NAD⁺, la fonction carboxylate du 3-PG et le groupement phosphate du 3-PG.

Parmi les composés identifiés nous retrouvons le tartrate qui est un ligand connu de l'enzyme. Nous avons aussi identifié le fluoropyruvate et la L-cystéine qui ont pu être co-cristallisés avec la forme tronquée de l'enzyme, ce qui a permis de montrer que ces ligands étaient capables de se lier au sein du site actif de l'enzyme. Finalement, la molécule **12** semble la plus intéressante de par sa capacité à se lier au sein du site de liaison du cofacteur et du substrat et l'interaction du groupement phényle dans la poche secondaire (Figure 7-2).

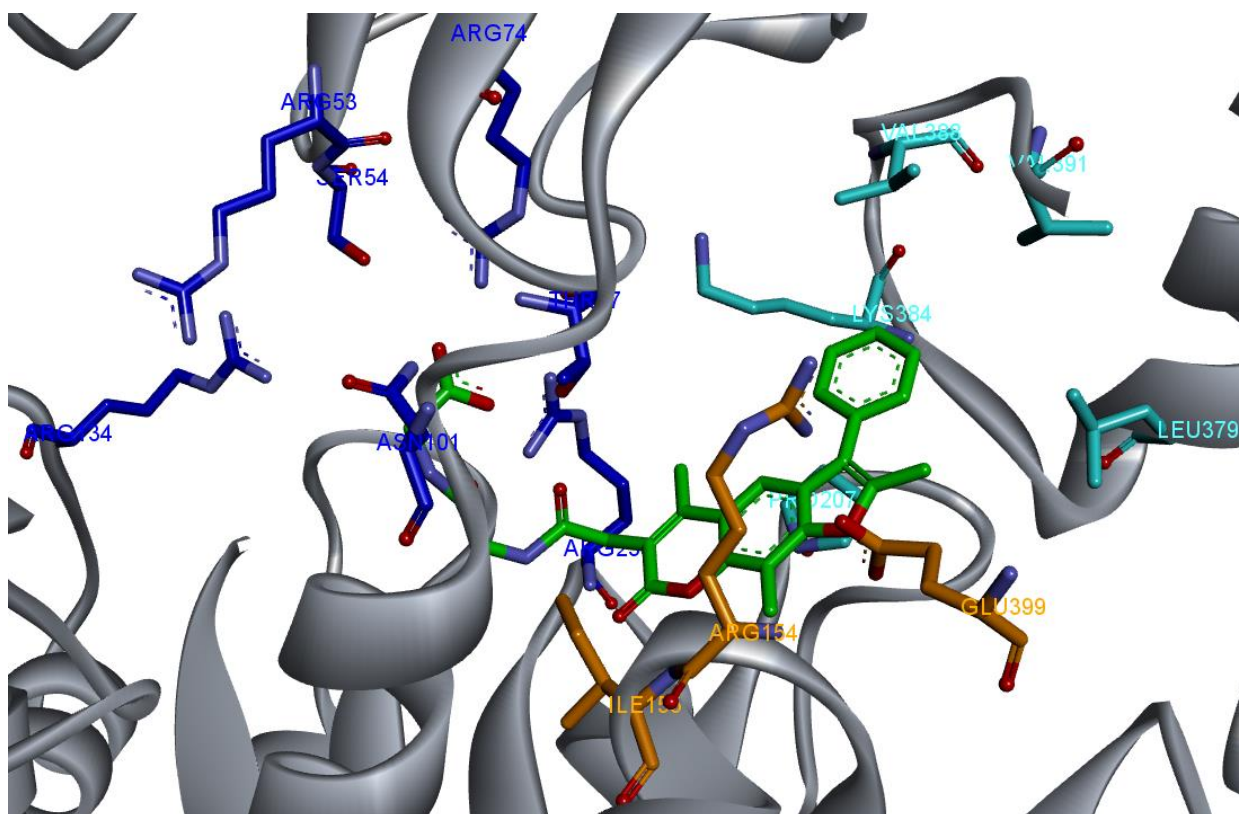


Figure 7-2: Site actif de la structure tétramérique de PHGDH modélisé en présence de la conformation du composé **12** (2-[[2-[[2-(7-keto-2,5,9-trimethyl-3-phenyl-furo[3,2-g]chromen-6-yl)acetyl]amino]acetyl]amino]acetate) simulée par docking (en vert). Les acides aminés mis en évidence lors du criblage virtuel sont représentés sous forme de bâtonnets. Les acides aminés de la poche du substrat sont en bleus, ceux de la poche du cofacteur en orange et ceux de la poche secondaire en cyan.

7.2 Perspectives

A l'heure actuelle la priorité est l'obtention de la structure de PHGDH sous sa forme complète et tétramérique. La cristallographie étant limitée par la réussite de l'étape de cristallisation, l'obtention d'une structure par cette méthode est compromise. La méthode de choix dans le cas de PHGDH (un complexe multimérique de grande taille) est la cryo-EM.

La compréhension de la perte d'activité totale de l'enzyme après mutation de la boucle flexible contenant R134 est la perspective la plus intéressante dans le cadre de ce projet. En plus de l'intérêt fondamental que constitue la compréhension de l'impact de la mutation de trois acides aminés qui ne semblent impliqués ni dans la liaison du substrat ni dans la stabilisation de l'état oligomérique de l'enzyme sur l'activité de celle-ci, cette

boucle pourrait représenter une cible pour le développement de nouveaux inhibiteurs de PHGDH (Figure 7-3).

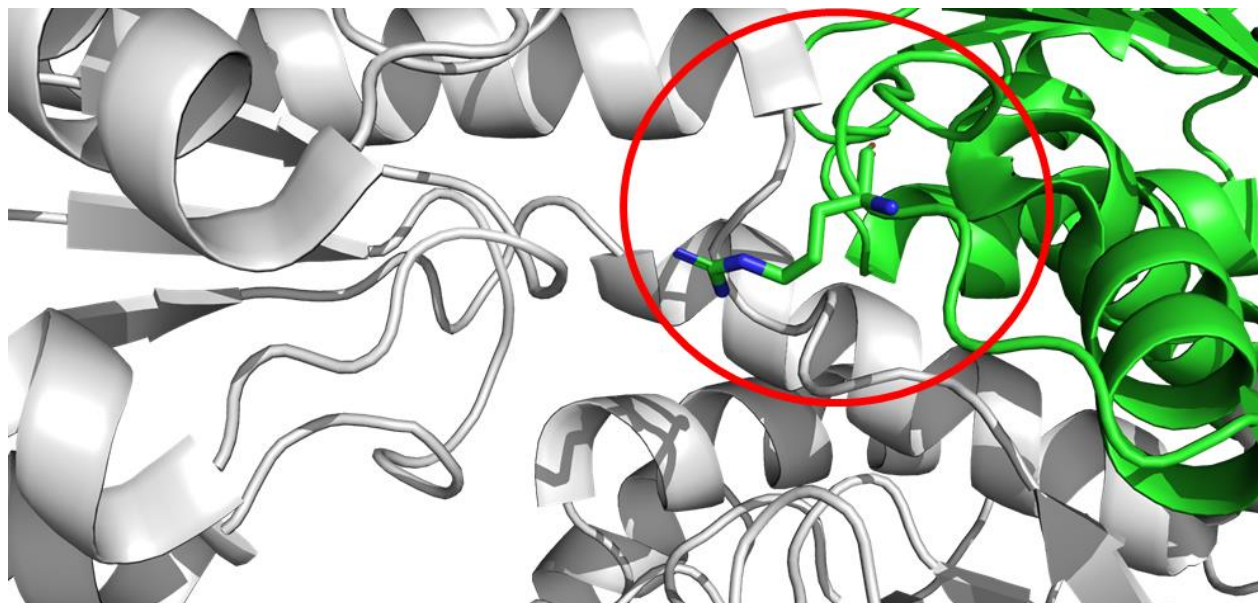


Figure 7-3: Interface entre deux monomères de PHGDH (en vert et en gris) au niveau du site actif. L'arginine 134 qui interagit avec le substrat est représenté sous forme de bâtonnet et la boucle flexible est mis en évidence.

Finalement, l'utilisation des résultats du criblage virtuel pour tester et concevoir de nouveaux inhibiteurs capables de cibler spécifiquement le site actif de PHGDH en formant des interactions dans la poche de liaison du cofacteur et du substrat en plus d'une poche secondaire est une perspective intéressante dans l'orientation thérapeutique du projet.

8. Matériel et méthodes

8 Matériel et méthodes

8.1.1 Thermoporation et expression des enzymes

Le protocole de thermoporation et d'expression étant commun à toutes les enzymes, il est présenté ici et a été utilisé pour toutes les constructions.

3µg de plasmide ont été insérés dans 50µL de bactéries *e. coli* BL21(DE3) (à saturation) thermo-compétentes par thermoporation (90s, 42°C). Les bactéries ont été incubées 1h à 37°C dans 1mL de milieu LB avant d'être étalées sur 2 plaques de LB agar comprenant l'antibiotique de résistance du plasmide. Les boîtes ont été incubées à 37°C durant une nuit. Une colonie a été sélectionnée et mise en préculture dans 10mL de LB (+ antibiotique) durant une nuit à 37°C. Les cultures (LB + antibiotique) ont été inoculées avec la préculture à un ratio de 40 :1. La culture a été incubée à 37° jusqu'à atteindre une $DO_{600} = 0,6$. Celle-ci est refroidie 30 minutes à 4°C avant d'induire la production de la protéine par ajout de 0,5mM IPTG durant 18 heures à 20°C.

Les bactéries sont finalement récupérées par centrifugation (JA-14, 5000g, 4°C, 15minutes).

8.1.2 PHGDH complète et mutation de surface

Ce protocole a été utilisé pour la purification de l'enzyme complète et des mutants de surface (réduction d'entropie de surface).

Les gènes codant pour les mutants de surface ont été commandés chez IDT⁹. La séquence de ceux-ci a été optimisée afin d'utiliser les codons les plus courants pour la bactérie *E. coli*. Ceux-ci ont été insérés dans le vecteur pET-28a (préalablement restreint en NdeI et XhoI) par ligation avec une enzyme T4 ligase durant la nuit à 20°C. La

⁹ <https://eu.idtdna.com/pages>

réaction de ligation a été effectuée dans un mélange composé de : 2µL buffer de ligation 5X, 1µL T4 DNA ligase, 1µL pET-28a (60ng/µL), 6µL de gène codant pour PHGDH mutée (30ng/µL). Le plasmide obtenu a été inséré (section 8.1.1) dans des bactéries Dh10b. Le plasmide des colonies identifiées comme positives à la présence du gène par réaction de polymérisation en chaîne (PCR) a été séquencé (primeur *forward* = TAGTTATTGCTCAGCGGTGG ; primeur *reverse* = AGCGGATAACAATTCCCCTC).

Le plasmide permettant l'expression de PHGDH nous a été envoyé par le Dr. Edouard Mullarky de l'université de Cornell (Etats-Unis). Ce plasmide est composé d'un vecteur pET-28a (résistant à la kanamycine) et du gène codant pour la forme sauvage de l'enzyme.

Les bactéries ont été resuspendues dans 60mL de tampon de lyse (50mM Tris (pH= 8,5), 10mM MgCl₂, 300mM NaCl, 10% glycérol, 5mM imidazole) par litre de culture en présence d'une pastille d'anti-protéase (Roche®, ref. Sigma : 11873580001). Une lyse par sonication est effectuée (10 bursts de 30s à 15W en alternance avec 30s de repos sur glace). Les débris bactériens sont séparés par centrifugation (JA-12, 10 000g, 4°C, 30minutes). Le surnageant est filtré (0,45µM et 0,2µM) et chargé sur une colonne IMAC (HisTrap FF crude, GE healthcare). La colonne est lavée avec le tampon de lavage (50mM Tris (pH= 8,5), 10mM MgCl₂, 300mM NaCl, 30mM imidazole, 10% glycérol) et la protéine est éluée avec le tampon d'élution (50mM Tris (pH= 8,5), 10mM MgCl₂, 250mM NaCl, 250mM imidazole, 10% glycérol). L'échantillon est dialysé durant la nuit à 4°C dans le tampon de dialyse (50mM Tris (pH= 8,5), 10mM MgCl₂, 250mM NaCl, 250mM imidazole, 10% glycérol). La solution protéique est ensuite chargée sur une colonne Superdex®200 (Ge Healthcare). La colonne est équilibrée avec le tampon d'équilibrage (50mM Tris (pH= 8,5), 10mM MgCl₂, 100mM NaCl, 10% glycérol) avant d'injecter la protéine et d'éluer celle-ci. PHGDH est finalement concentrée à la concentration de cristallisation et conservée dans 50mM Tris (pH= 8,5), 10mM MgCl₂, 100mM NaCl, 20% glycérol à -80°C.

8.1.3 PHGDH fusionnée à une MBP

Le plasmide permettant l'expression de PHGDH fusionnée à une *Matose Binding Protein* a été fourni nos collaborateurs (laboratoire du professeur Olivier Féron) de l'université catholique de Louvain (UCL, Belgique). Ce plasmide est composé d'un vecteur pVP16 (résistant à l'ampicilline) et du gène codant pour la forme sauvage de l'enzyme fusionnée à une MBP.

Les bactéries ont été resuspendues dans 15mL de tampon de lyse (50mM Tris (pH= 7,5), 150mM NaCl, 1% triton x-100) par gramme de culot bactérien en présence d'une pastille d'anti-protéase (Roche®, ref. Sigma : 11873580001). Une lyse par sonication est effectuée (10 bursts de 30s à 15W en alternance avec 30s de repos sur glace). Les débris bactériens sont séparés par centrifugation (JA-12, 10 000g, 4°C, 30minutes). Le surnageant est filtré (0,45µM et 0,2µM) et chargé sur une colonne IMAC (HisTrap FF crude, GE healthcare). Une colonne MBPTrap (GE healthcare) est connectée en série. La colonne est lavée avec le tampon de lavage 1 (10mM Tris (pH= 7,5), 150mM NaCl) et la protéine est éluée de la colonne IMAC avec le tampon d'élution (50mM Tris (pH= 7,5), 100mM NaCl, 500mM imidazole). Le tampon de lavage 2 (50mM Tris (pH= 7,5), 100mM NaCl, 200mM imidazole, 2mM maltose) permet d'éluer l'enzyme de la colonne MBPTrap. PHGDH est finalement concentrée à la concentration de cristallisation et conservée dans 50mM Tris (pH= 7,5), 100mM NaCl, 200mM imidazole, 2mM maltose, 20% glycérol à -80°C.

8.1.4 sPHGDH (3-338)

Le plasmide permettant l'expression de sPHGDH (3-338) a été offert par Nicola Burgess-Brown (Addgene plasmid # 38841 ; <http://n2t.net/addgene:38841> ; RRID:Addgene_38841). Ce plasmide est composé d'un vecteur pNIC28-Bsa4 (résistant à la kanamycine) et du gène codant pour la forme tronquée de l'enzyme.

Le mutant R134A a été obtenu par PCR successives. L'ajout de la mutation a été effectué à l'aide des primeurs représentés au Tableau 8-1. Les PCR 1 et 2 ont permis d'induire la mutation sur les deux segments du gène. Une troisième PCR a permis de synthétiser le gène complet codant pour sPHGDH avec la mutation. Le protocole de PCR est identique pour les 3 PCR (4µL dNTPs, 4µL primeur *forward* (10µM), 4µL primer *reverse* (10µM), 25µL tampon Takara 2x, 0,5µL plasmid, 12µL dH₂O, 0,5µL Primestart® polymérase (Ozyme, TAKR050A). Les paramètres utilisés lors de la PCR sont repris au Tableau 8-2.

Tableau 8-1: Séquence des primeurs utilisés pour insérer la mutation R134A dans l'enzyme sPHGDH. Le codon correspondant à la mutation est souligné.

	Primeur <i>forward</i>	Primeur <i>reverse</i>
PC R 1	CATATGGCAAATCTGCGGAAAGT GCTC	CATGAACTTCTT <u>CCG</u> CTCCCAT TTGC
PC R 2	GCAAATGGGAG <u>CGG</u> AAGAAGTT CATG	CTCGAGACCATCGCTTCTTTTCG GGC

Tableau 8-2: Paramètres utilisés pour la PCR en 3 étapes.

Cycle	Temperature	Time
1	98°C	30s
30	98°C	10sec
	58°C	10sec
	72°C	30sec/500bp
1	72	120s

Le gène obtenu est alors restreint avec les enzymes NdeI et XhoI et la ligation est effectuée suivant le même protocole que celui présenté au point 3.1.2.

Dans le cadre de sPHGDH, les bactéries utilisées sont des Rosetta (DE3).

Les bactéries ont été resuspendues dans 10mL de tampon de lyse (50mM NaH₂PO₄ (pH 8,0), 300mM NaCl, 30mM imidazole, 0,5mM TCEP) par gramme de culot bactérien. 0,2 mg/mL de lysozyme, 25U benzonase (Sigma) pour 10mL de tampon de lyse et une

tablette d'antiprotéase EDTA-free (Roche) ont été ajoutés et l'échantillon a été incubé 30 minutes à température ambiante. Une lyse par sonication est effectuée (10 bursts de 10s à 30W en alternance avec 30s de repos sur glace). Les débris bactériens sont séparés par centrifugation (JA-12, 12 000g, 4°C, 25 minutes). Le surnageant est filtré (0,45µM et 0,2µM) et chargé sur une colonne IMAC (HisTrap FF crude, GE healthcare). La colonne est préalablement équilibrée avec le tampon de lavage 1 (50mM NaH₂PO₄ (pH 8,0), 300mM NaCl, 30mM imidazole, 0,5mM TCEP). La protéine est éluée dans le tampon d'élution 1 (50mM NaH₂PO₄ (pH 8,0), 300mM NaCl, 300mM imidazole, 0,5mM TCEP). L'échantillon est ensuite dialysé durant la nuit à 4°C dans le tampon de dialyse 1 (50mM NaH₂PO₄ (8,0), 300mM NaCl, 30mM imidazole, 0,25mM TCEP). L'His₆-tag de la protéine est clivé avec l'ajout de rTEV protéase à un ratio molaire de 1 :20 et incubé à température ambiante durant 3 heures. L'échantillon est chargé sur une colonne IMAC (équilibrée avec le tampon de lavage). La protéine clivée est éluée avec le tampon de lavage 1 alors que l'His₆-tag et la protéine non clivée sont élués avec le tampon d'élution 1.

L'enzyme clivée est dialysée durant la nuit à 4°C dans le tampon de dialyse 2 (20mM Tris (7,5), 40mM NaCl, 0,25mM TCEP). L'échantillon est chargé sur une colonne Q HP (GE Healthcare) préalablement équilibrée avec le tampon de lavage 2 (20mM Tris (7,5), 40mM NaCl, 0,5mM TCEP) et éluée avec un gradient de 0 à 60% en tampon élution 2 (20mM Tris (7,5), 1M NaCl, 0,5mM TCEP) réalisé en 50 minutes. A l'aide d'une colonne PD-10 (Ge Healthcare), sPHGDH est transférée dans le tampon de conservation (25mM HEPES (7,5), 100mM NaCl, 0,5mM TCEP) et conservée à -80°C à une concentration comprise entre 2 et 3 mg/mL.

8.1.5 Essais de cristallisation

Les enzymes PHGDH complète, MBP-PHGDH et les 3 mutants de surface étaient conservés à la concentration de cristallisation. sPHGDH a été concentré après décongélation à la concentration de cristallisation. Lors des essais de co-cristallisation,

le ligand a été ajouté à la protéine à un ratio molaire de 1 : 5 (protéine : ligand) à la protéine non concentrée et a été incubé 2 heures à température ambiante.

La protéine concentrée a été centrifugée à 12000g durant 10minutes avant utilisation.

Le criblage de conditions avec les kits commerciaux (Hampton Research 1 et 2 (HR2-110 ; HR2-112) et Wizard CRYO 1 et 2 (1009536 ; 1009537)) a été réalisé en plaque 96 puits (goutte assise, $V_{\text{réservoir}} = 50 \mu\text{L}$, $V_{\text{enzyme}} = 1 \mu\text{L}$) en ratio 1 : 1 et 2 : 1 (réservoir : enzyme). Toutes ces conditions ont été réalisées en présence et en absence de NAD^+ .

Les essais cristallisations ont été réalisé dans des plaques 24 puits (goutte pendante, $V_{\text{réservoir}} = 700 \mu\text{L}$, $V_{\text{enzyme}} = 1 \mu\text{L}$). Toutes les solutions ont été préparées et filtrées avant utilisation pour la cristallisation. Tous les essais de cristallisation ont été fait manuellement.

Les boites de cristallisation ont été scellées et conservées à température ambiante (~19-21°C). Celles-ci étaient examinées les premières semaines et ensuite une fois par mois afin de suivre la formation des cristaux.

8.1.6 Collecte et traitement des données cristallographiques

Les cristaux les plus prometteurs ont été analysés au synchrotron SOLEIL (Gif-sur-Yvette, France) sur les lignes PROXIMA 1 et PROXIMA 2.

Le traitement des données obtenues a été réalisé par le Dr. Serena Sirigu. L'intégration des données a été réalisée à l'aide du programme XDS. Suite à la haute anisotropie des données, le programme STARANISO (Global Phasing) a été utilisé pour générer un nouveau jeu de données en formant un ellipsoïde orienté dans les directions de l'espace présentant le plus de réflexions. La génération d'un modèle et l'affinement de celui-ci a été réalisé avec BUSTER (Global Phasing). L'affinement manuel et le placement des ligands a été réalisé avec WinCOOT (0.8.9). Les données sont visualisées grâce à PyMOL (2.2.2), Discovery studio visualizer 2016 et WinCOOT (0.8.9).

8.1.7 Tests enzymatiques

Les tests réalisés par notre collaborateur le Dr E. Mullarky ont été faits dans les mêmes conditions que celles déjà publiées par celui-ci^{59,67,68}.

Les tests que nous avons réalisés ont été faits dans un mix réactionnel composé de : 100mM Tris (pH = 8,8), 400 mM NaCl, 1 mM NAD⁺, 1,4 µg/50 µL PHGDH. Le 3-phosphoglycérate (concentration variable) permettait d'initier la réaction. Les mesures de fluorimétrie étaient réalisées à une longueur d'onde d'excitation = 360 nm et d'émission = 460 nm (Beckmann Coulter DTX 880 multimode detector). Une mesure était faite toutes les 10 secondes durant 10 minutes.

8.1.8 Génération du modèle

Deux modèles ont été générés indépendamment avec le programme SWISS-MODEL¹⁰ auquel nous avons fourni la structure de la forme bactérienne de PHGDH provenant de *Mycobacterium tuberculosis* et avec le programme I-TASSER¹¹. La séquence de l'enzyme a été fournie aux deux programmes.

8.1.9 Couplage des états d'oligomérisation

Le mélange réactionnel est composé de 2mM NAD⁺, 10mM HEPES et 1 µg PHGDH / 6µL mix et est incubé 30 minutes sur glace. Le couplage est initié par ajout de 5 mM BS3 et la solution est incubée à température ambiante sous agitation durant 30 minutes. La réaction est arrêtée avec 0,5M Tris (pH = 7,5) et l'incubation est prolongée pour 15 minutes supplémentaires. L'échantillon est conservé à -80°C avant d'être digéré.

¹⁰ <https://swissmodel.expasy.org/>

¹¹ <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>

8.1.10 Digestion de l'enzyme et acquisition des données par spectrométrie de masse.

La digestion de l'enzyme a été réalisée en suivant le protocole décrit par *Distler et al.*¹⁰² Les enzymes de restriction utilisées sont la trypsine, la chymotrypsine, la LysC et un mélange trypsine/chymotrypsine.

L'acquisition des données par spectrométrie de masse a été réalisée par Marc Dieu (URBC, UNamur). Les échantillons ont été analysés par LC-MS-MS (nanoRSLC Ultimate 3000, Thermofischer, colonne : PEPMAP 100 (Thermo) de 25cm x 75µm; maXis Impact Q-TOF, Bruker). La LC a appliqué un gradient de 4 à 35% en tampon B durant 35 minutes ensuite augmenté jusqu'à 90% en tampon B en 15 min, pour finalement redescendre jusqu'à 4% en tampon B durant 20 minutes (tampon A : 2% ACN, 98% H₂O, 0.1% acide formique ; tampon B : 80% ACN, 20% H₂O, 0.1% acide formique).

Les données ont été acquises durant 0,5 sec pour une gamme de m/z de 50 à 2200. Ensuite, pendant 2,5 sec, les peptides les plus abondants de charges 2+, 3+ ou 4+ sont séquencés. L'énergie dans la cellule de collision (CID) est fonction du rapport m/z et de l'état de charge du peptide. Les ions séquencés sont conservés si leur abondance augmente d'un facteur 3 en 20 secondes. Le logiciel Compass HyStar 3.2 (Bruker) est utilisé sur le spectromètre de masse.

Le logiciel DataAnalysis 4.2 (Bruker) a permis de créer la liste des pics qui comprend les masses des peptides séquencés (les ions précurseurs ou les ions mères) et les masses des fragments (ions filles).

Les données obtenues ont été exportées sur Stavrox (3.6.6) qui a permis la mise en évidence des couplages potentiels. L'analyse des données et la visualisation des distances a été faite manuellement à l'aide du logiciel PyMOL (2.2.2). La visualisation des peptides couplés a été réalisée avec xiNET^{12,103}.

¹² <http://crosslinkviewer.org/>

8.1.11 Génération de la position de ligands au sein du site actif

Les molécules ont été dessinées avec Chemdraw (18.1) et sauvegardées en .mol2 ou téléchargées en .mol2 via la base de données ZINC^{91,13}. La conformation des ligands a été optimisée par minimisation de leur énergie avec le champ de force MMFF (Merck Molecular Force Field) (algorithme : gradient conjugué ; nombre cycle maximum 2000 ; gradient RMS = 0,01 kcal/molÅ). La protéine a été préparée à l'aide du champ de force CHARMM (Chemistry at HARvard Macromolecular Mechanics).

La génération de conformations de ligands au sein du site actif a été réalisée avec le logiciel Gold (2.3.3) sur les installations du Consortium des Équipements de Calculs Intensifs (CÉCI) et plus particulièrement sur la Plateforme Technologique de Calculs Intensifs (PTCI) à l'Université de Namur. Pour chaque ligand, le calcul a été réalisé dans une sphère de 10 Å centrée sur le centre de masse de ligand cristallisé au sein de la poche (cofacteur ou substrat). 20 conformations ont été générées dont le score a été déterminé par la fonction de score Chemscore. La conformation présentant le plus haut score a été sélectionnée.

Les conformations et les modes de liaisons des ligands ont été visualisés avec Discovery Studio visualizer 2016 (pont hydrogène > 3,4 Å, 90 à 180° ; interaction halogène > 3,7 Å, 120 à 180° ; π -cation > 5 Å, 0 à 40° ; $\pi - \pi$ > 6 Å, 0 à 50° ; interaction alkyle > 5,5 Å).

¹³ <http://zinc.docking.org/>

9. Références

9 Références

1. Cancer. Available at: <https://www.who.int/news-room/fact-sheets/detail/cancer>. Consulté le 02/10/2019.
2. Le cancer en chiffres. Available at: <https://kankerregister.org/default.aspx?lang=FR>. Consulté le 02/10/2019.
3. Degos L. John Hughes Bennett, Rudolph Virchow... and Alfred Donné: the first description of leukemia. *Hematol. J. Off. J. Eur. Haematol. Assoc.* **2**, 1 (2001).
4. National Cancer Act of 1971. Available at: https://dtp.cancer.gov/timeline/noflash/milestones/M4_Nixon.htm. Consulté le 02/10/2019.
5. Evolution and Future of Cancer Treatments. <https://medium.com/what-will-it-take-to-end-cancer/evolution-and-future-of-cancer-treatments-7241c4a005a5>. Consulté le 02/10/2019.
6. What Is Cancer? Available at: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>. Consulté le 02/10/2019.
7. Forbes, S. A. *et al.* COSMIC: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43**, D805–D811 (2015).
8. Vanneman, M. & Dranoff, G. Combining immunotherapy and targeted therapies in cancer treatment. *Nat. Rev. Cancer* **12**, 237–251 (2012).
9. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100**, 57–70 (2000).
10. Hanahan, D. & Weinberg, R. a. Hallmarks of cancer: the next generation. *Cell* **144**, 646–74 (2011).
11. Otto Warburg. On the origin of cancer cells. *Science (80-.)*. **123**, 309–14 (1956).
12. Warburg, O. On respiratory impairment in cancer cells. *Science (80-.)*. **124**, 269–70 (1956).

13. The Nobel Prize in Physiology or Medicine 1931. Available at: <https://www.nobelprize.org/prizes/medicine/1931/summary/>. Consulté le 02/10/2019.
14. Racker, E. Bioenergetics and the problem of tumor growth. *Am. Sci.* **60**, 56–63 (1972).
15. Boerner, P., Resnick, R. J. & Racker, E. Stimulation of glycolysis and amino acid uptake in NRK-49F cells by transforming growth factor beta and epidermal growth factor. *Proc. Natl. Acad. Sci.* **82**, 1350–1353 (1985).
16. Flier, J. S., Mueckler, M. M., Usher, P. & Lodish, H. F. Elevated levels of glucose transport and transporter messenger RNA are induced by ras or src oncogenes. *Science (80-.).* **235**, 1492–1495 (1987).
17. Birnbaum, M. J., Haspel, H. C. & Rosen, O. M. Transformation of Rat Fibroblasts by FSV Rapidly Increases Glucose Transporter Gene Transcription. *Science (80-.).* **235**, 1495–1498 (1987).
18. Y., H., Rosen, O. M. & Birnbaum, M. J. Growth factors rapidly induce expression of the glucose transporter gene. *J. Biol. Chem.* **263**, 13655–13662 (1988).
19. Shim, H., Chun, Y. S., Lewis, B. C. & Dang, C. V. A unique glucose-dependent apoptotic pathway induced by c-Myc. *Proc. Natl. Acad. Sci.* **95**, 1511–1516 (1998).
20. Fantin, V. R., St-Pierre, J. & Leder, P. Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. *Cancer Cell* **9**, 425–434 (2006).
21. Liberti, M. V & Locasale, J. W. The Warburg Effect : How Does it Bene fi t Cancer Cells ? *Trends Biochem. Sci.* **41**, 211–218 (2016).
22. Vander Heiden, M. G., Cantley, L. C. & Thompson C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science (80-.).*

-). **324**, 1029–1033 (2009).
23. Locasale, J. W. & Cantley, L. C. Metabolic flux and the regulation of mammalian cell growth. *Cell Metab.* **14**, 443–451 (2011).
 24. Shestov, A. A. *et al.* Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. *Elife* **3**, 1–18 (2014).
 25. Pfeiffer, T. Cooperation and Competition in the Evolution of ATP-Producing Pathways. **504**, (2014).
 26. Slavov, N., Budnik, B. A., Schwab, D., Airoidi, E. M. & van Oudenaarden, A. Constant Growth Rate Can Be Supported by Decreasing Energy Flux and Increasing Aerobic Glycolysis Nikolai. *Cell Rep.* **7**, 705–714 (2014).
 27. Chang, C. *et al.* Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell* **162**, 1229–1241 (2015).
 28. Ho, P. *et al.* Phosphoenolpyruvate Is a Metabolic Checkpoint of Anti-tumor T Cell Responses. *Cell* **162**, 1217–1228 (2015).
 29. Estrella, V., Chen, T., Lloyd, M. & Wojtkowiak, J. *et al.* Acidity generated by the tumor microenvironment drives local invasion. *Cancer Res.* **73**, 1524–1535 (2013).
 30. Robert A. Gatenby, E. T. G. A Reaction-Diffusion Model of Cancer Invasion. *Cancer Res.* **56**, 5745–5753 (1996).
 31. Oscar R. Colegio, Ngoc-Quynh Chu, A. L. S. *et al.* Functional polarization of tumour-associated macrophages by tumour-derived lactic acid Oscar. *Nature* **513**, 559–563 (2014).
 32. Molenaar, D., Berlo, R. Van & Ridder, D. De. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 1–10 (2009).
 33. Locasale, J. W. The consequences of enhanced cell-autonomous glucose metabolism. *Trends Endocrinol. Metab.* **23**, 545–551 (2012).

34. Kaplon, J. *et al.* A key role for mitochondrial gatekeeper pyruvate dehydrogenase in oncogene-induced senescence. *Nature* **498**, 109–112 (2013).
35. Levine, A. J. & Puzio-kuter, A. M. The Control of the Metabolic Switch in Cancers by Oncogenes and Tumor Suppressor Genes. *Science* (80-.). **330**, 1340–1344 (2010).
36. Dang, C. V. Links between metabolism and cancer. *Genes Dev.* **26**, 877–890 (2012).
37. Koppenol, W. H., Bounds, P. L. & Dang, C. V. Otto Warburg's contributions to current concepts of cancer metabolism. *Nat. Publ. Gr.* **11**, 325–337 (2011).
38. Boroughs, L. K. & Deberardinis, R. J. Metabolic pathways promoting cancer cell survival and growth. *Nat. Publ. Gr.* **17**, 351–359 (2015).
39. Fell, D. A. & Snell, K. Control analysis of mammalian serine biosynthesis. Feedback inhibition on the final step. *Biochem. J.* **256**, 97–101 (1988).
40. Snell, K., Natsumeda, Y., Eble, J. N., Glover, J. L. & Weber, G. Enzymic imbalance in serine metabolism in human colon carcinoma and rat sarcoma. *Br. J. Cancer* **57**, 87–90 (1988).
41. Furuya, S. An essential role for de novo biosynthesis of L-serine in CNS development. *Asia Pac J Clin Nutr* **17**, 312–315 (2008).
42. Kalhan, S. C. & Hanson, R. W. Resurgence of serine: An often neglected but indispensable amino acid. *J. Biol. Chem.* **287**, 19786–19791 (2012).
43. Tom de Koning, L. K. Serine-deficiency syndromes. *Curr. Opin. Neurol.* **17**, 197–204 (2004).
44. Yoshida, K. *et al.* Targeted Disruption of the Mouse 3-Phosphoglycerate Dehydrogenase Gene Causes Severe Neurodevelopmental Defects and Results in Embryonic Lethality. *J. Biol. Chem.* **279**, 3573–3577 (2004).
45. Jaeken, J. *et al.* 3-Phosphoglycerate dehydrogenase deficiency and 3-

- phosphoserine phosphatase deficiency: Inborn errors of serine biosynthesis. *J. Inherit. Metab. Dis.* **19**, 223–226 (1996).
46. Hart, C. E. *et al.* Phosphoserine Aminotransferase Deficiency: A Novel Disorder of the Serine Biosynthesis Pathway. *Am. J. Hum. Genet.* **80**, 931–937 (2007).
 47. Shan, C. *et al.* Phosphoglycerate mutase 1 coordinates glycolysis and biosynthesis to promote tumor growth. *Cancer Cell* **22**, 585–600 (2012).
 48. Hay, N. Reprogramming glucose metabolism in cancer: can it be exploited for cancer therapy? *Nat. Rev. Cancer* **16**, 1–15 (2016).
 49. Lewis, C. A. *et al.* Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. *Mol. Cell* **55**, 253–263 (2014).
 50. Fan, J. *et al.* Quantitative flux analysis reveals folate-dependent NADPH production. *Nature* **510**, 298–302 (2014).
 51. Pacold, M. E. *et al.* A PHGDH inhibitor reveals coordination of serine synthesis and one-carbon unit fate. *Nat. Chem. Biol.* (2016). doi:10.1038/nchembio.2070
 52. Locasale, J. W. *et al.* Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat. Genet.* **43**, 869–874 (2011).
 53. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476**, 346–350 (2011).
 54. Grant, G. A. Contrasting catalytic and allosteric mechanisms for phosphoglycerate dehydrogenases. *Arch. Biochem. Biophys.* **519**, 175–185 (2012).
 55. Unterlass, J. E. *et al.* Structural insights into the enzymatic activity and potential substrate promiscuity of human 3-phosphoglycerate dehydrogenase (PHGDH). *Oncotarget* **8**, 104478–104491 (2017).
 56. O43175. Available at: <https://www.uniprot.org/uniprot/O43175>.

57. Aravind, L. & Koonin, E. V. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *J. Mol. Biol.* **287**, 1023–1040 (1999).
58. Dey, S., Grant, G. A. & Sacchettini, J. C. Crystal Structure of Mycobacterium tuberculosis D-3-Phosphoglycerate Dehydrogenase: EXTREME ASYMMETRY IN A TETRAMER OF IDENTICAL SUBUNITS. *J. Biol. Chem.* **280**, 14892–14899 (2005).
59. Mullarky, E. *et al.* Inhibition of 3-phosphoglycerate dehydrogenase (PHGDH) by indole amides abrogates de novo serine synthesis in cancer cells. *Bioorg. Med. Chem. Lett.* (2019). doi:10.1016/j.bmcl.2019.07.011
60. Unterlass, J. E. *et al.* Validating and enabling phosphoglycerate dehydrogenase (PHGDH) as a target for fragment-based drug discovery in PHGDH-amplified breast cancer. *Oncotarget* **5**, (2016).
61. Dey, S., Burton, R. L., Grant, G. A. & Sacchettini, J. C. Structural Analysis of Substrate and Effector Binding in Mycobacterium tuberculosis D-3-Phosphoglycerate Dehydrogenase. *Biochemistry* **47**, 8271–8282 (2008).
62. Dey, S., Hu, Z., Xiao, L. X., Sacchettini, J. C. & Grant, G. A. The effect of hinge mutations on effector binding and domain rotation in Escherichia coli D-3-phosphoglycerate dehydrogenase. *J. Biol. Chem.* **282**, 18418–18426 (2007).
63. Thompson, J. R., Bell, J. K., Bratt, J., Grant, G. A. & Banaszak, L. J. Vmax regulation through domain and subunit changes. The active form of phosphoglycerate dehydrogenase. *Biochemistry* **44**, 5763–5773 (2005).
64. Bell, J. K., Grant, G. a & Banaszak, L. J. Multiconformational states in phosphoglycerate dehydrogenase. *Biochemistry* **43**, 3450–8 (2004).
65. D.J. Schuller, G.A. Grant, L. J. B. The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase. *Nat.Struct.Mol.Biol* **2**, 69–76 (1995).

66. Singh, R. K., Raj, I., Pujari, R. & Gourinath, S. Crystal structures and kinetics of Type III 3-phosphoglycerate dehydrogenase reveal catalysis by lysine. *FEBS J.* **281**, 5498–5512 (2014).
67. Mullarky, E. *et al.* Identification of a small molecule inhibitor of 3-phosphoglycerate dehydrogenase to target serine biosynthesis in cancers. *Proc. Natl. Acad. Sci.* **113**, 201521548 (2016).
68. Ravez, S. *et al.* α -Ketothioamide Derivatives: A Promising Tool to Interrogate Phosphoglycerate Dehydrogenase (PHGDH). *J. Med. Chem.* [acs.jmedchem.6b01166](https://doi.org/10.1021/acs.jmedchem.6b01166) (2017). doi:10.1021/acs.jmedchem.6b01166
69. Spillier, Q. *et al.* Anti-alcohol abuse drug disulfiram inhibits human PHGDH via disruption of its active tetrameric form through a specific cysteine oxidation. *Sci. Rep.* **9**, 4737 (2019).
70. Fuller, N. *et al.* An improved model for fragment-based lead generation at AstraZeneca. *Drug Discov. Today* (2016). doi:10.1016/j.drudis.2016.04.023
71. Smyth, D. & Mrozkiewicz, M. Crystal structures of fusion proteins with large affinity tags. *Protein Sci.* 1313–1322 (2003). doi:10.1110/ps.0243403.the
72. Waugh, D. S. Crystal structures of MBP fusion proteins. *Protein Sci.* **25**, 559–571 (2016).
73. Derewenda, Z. S. Rational protein crystallization by mutational surface engineering. *Structure* **12**, 529–35 (2004).
74. Derewenda, Z. S. It's all in the crystals... *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 243–248 (2011).
75. Madrona, Y. *et al.* Crystal Structure of Cindoxin, the P450cin Redox Partner. *Biochemistry* **53**, 1435–1446 (2014).
76. Doig, J. & Sternberg, M. J. Side-chain conformational entropy in protein folding. *Protein Sci.* **4**, 2247–51 (1995).

77. Rodrigues, J. V, Prosinecki, V., Marrucho, I., Rebelo, L. P. N. & Gomes, C. M. Protein stability in an ionic liquid milieu: on the use of differential scanning fluorimetry. *Phys. Chem. Chem. Phys.* **13**, 13614–6 (2011).
78. Sorrell, F. J., Greenwood, G. K., Birchall, K. & Chen, B. Development of a differential scanning fluorimetry based high throughput screening assay for the discovery of affinity binders against an anthrax protein. *J. Pharm. Biomed. Anal.* **52**, 802–808 (2010).
79. Czepas, J. & Krowarsch, D. The impact of Lys Arg surface mutations on the crystallization of the globular domain of RhoGDI. *Acta Cryst. D* **60**, 275–280 (2004).
80. Xu, X. L. & Grant, G. A. Regulation of mycobacterium tuberculosis d -3-phosphoglycerate dehydrogenase by phosphate-modulated quaternary structure dynamics and a potential role for polyphosphate in enzyme regulation. *Biochemistry* **53**, 4239–4249 (2014).
81. Biasini, M. *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, 252–258 (2014).
82. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
83. Benkert, P., Biasini, M. & Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**, 343–350 (2011).
84. Walzthoeni, T., Leitner, A., Stengel, F. & Aebersold, R. Mass spectrometry supported determination of protein complex structure. *Curr. Opin. Struct. Biol.* **23**, 1–9 (2013).
85. Rappsilber, J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540 (2011).

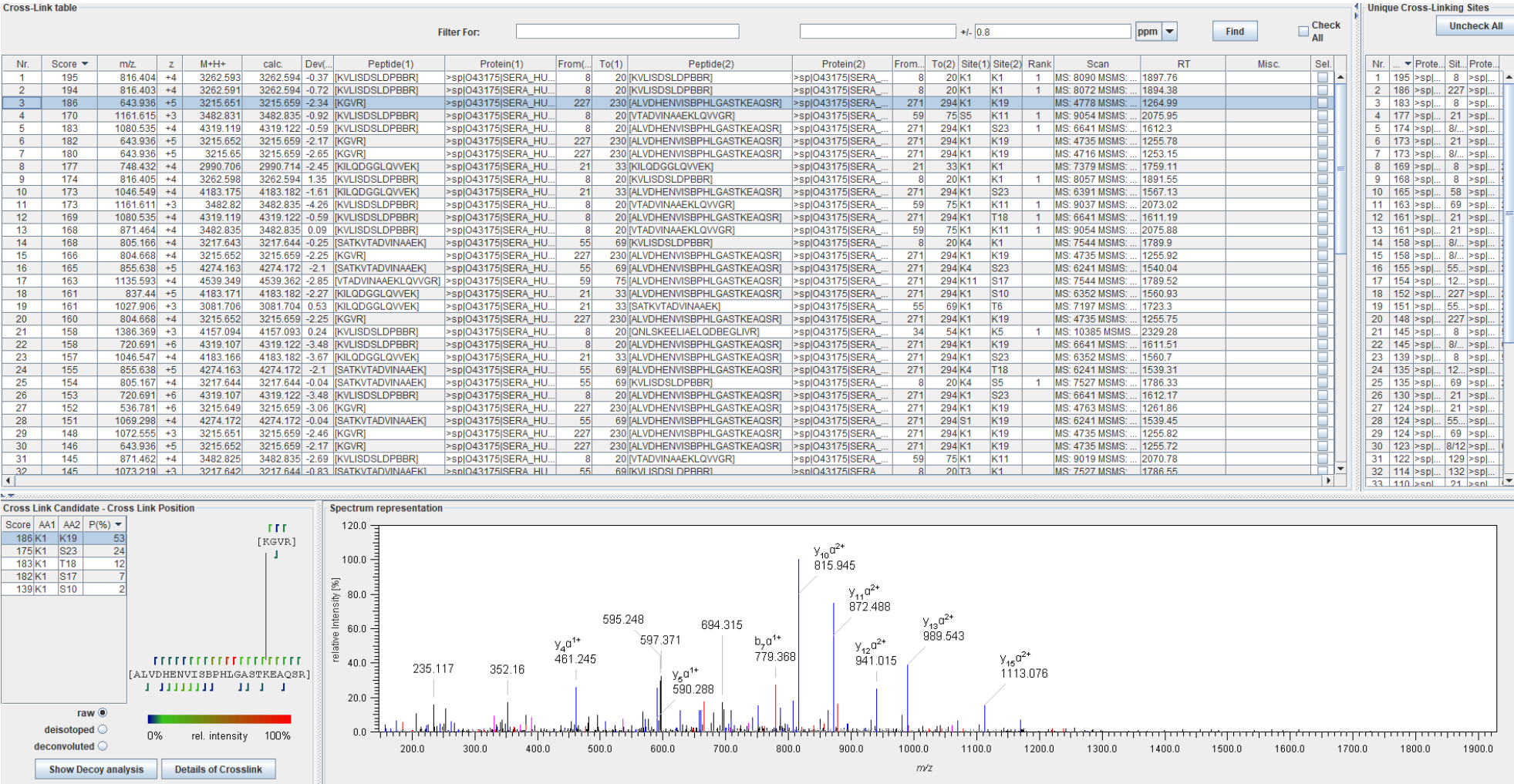
86. Leitner, A. *et al.* Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics. *Mol. Cell. Proteomics* **9**, 1634–1649 (2010).
87. Götze, M. *et al.* StavroX-A software for analyzing crosslinked products in protein interaction studies. *J. Am. Soc. Mass Spectrom.* **23**, 76–87 (2012).
88. Feig, S. A., Segel, G. B., Shohet, S. B. & Nathan, D. G. Energy metabolism in human erythrocytes. *J. Clin. Invest.* **50**, 1731–1737 (1971).
89. Shigeki, M., Takayasu, S., Chiyo, S. & Haruhisa, Y. The hydrogen ion concentrations and erythrocyte glycolysis. *Biochem. Biophys. Res. Commun.* **17**, 748–751 (1964).
90. Mulquiney, P. J. & Kuchel, P. W. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations1: equations and parameter refinement. *Biochem. J.* **342**, 581–596 (1999).
91. Sterling, T. & Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
92. K. D. Bird. The use of spermicide containing nonoxynol-9 in the prevention of HIV infection. *Aids* **5**, 791–796 (1991).
93. J. Stephenson. Vaccines pose no diabetes, bowel disease risk. *Jama* **284**, 2307–2308 (2000).
94. Knorr, B. *et al.* Montelukast for chronic asthma in 6- to 14-year-old children: A randomized, double-blind trial. *J. Am. Med. Assoc.* **279**, 1181–1186 (1998).
95. Markham, A. & Faulds, D. Montelukast. *Drugs* **56**, 251–256 (1998).
96. Osborne, C. K., Wakeling, A. & Nicholson, R. I. Fulvestrant: An oestrogen receptor antagonist with a novel mechanism of action. *Br. J. Cancer* **90**, S2–S6 (2004).
97. David Murdoch, D. M. Fosinopril: A Review of its Pharmacodynamic and Pharmacokinetic Properties, and Therapeutic Potential in Hypertension. *Drugs*

- 43**, 123–140 (1992).
98. Becker, B. & Gage, T. Demecarium Bromide and Echothiophate Iodide in Chronic Glaucoma. *AMA. Arch. Ophthalmol.* **63**, 102–107 (1960).
 99. Adjei, A. A. Pemetrexed (Alimta®): A novel multitargeted antifolate agent. *Expert Rev. Anticancer Ther.* **3**, 145–156 (2003).
 100. Tarpley, T. Lapatinib in the treatment of breast cancer. *Expert Rev. Anticancer Ther.* **7**, 1183–1192 (2007).
 101. Frueh, D. P., Goodrich, A., Mishra, S. & Nichols, S. NMR methods for structural studies of large monomeric and multimeric proteins. *Curr. Opin. Struct. Biol.* **23**, 734–739 (2013).
 102. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
 103. Combe, C. W., Fischer, L. & Rappsilber, J. xiNET: Cross-link Network Maps With Residue Resolution. *Mol. Cell. Proteomics* **14**, 1137–1147 (2015).

10. Annexes

10 Annexes

Annexe A : Exemple de résultats obtenus pour les potentiels couplages visualisé avec StavroX.



Annexe B : Exemple d'input utilisé pour un calcul de docking avec GOLD.

GOLD CONFIGURATION FILE

AUTOMATIC SETTINGS

autoscale = 2

POPULATION

popsiz = auto

select_pressure = auto

n_islands = auto

maxops = auto

niche_siz = auto

GENETIC OPERATORS

pt_crosswt = auto

allele_mutatewt = auto

migratewt = auto

FLOOD FILL

radius = 10

origin = 50.732 104.52 54.8273

do_cavity = 1

floodfill_atom_no = 0

cavity_file =

floodfill_center = point

DATA FILES

ligand_data_file

/home/qthemans/Documents/phgdh/Docking/Ppyr_fpyr/molecules.mol2 20

param_file = DEFAULT

set_ligand_atom_types = 1

set_protein_atom_types = 0

directory = /home/qthemans/Documents/phgdh/Docking/Ppyr_fpyr/MonoA

tordist_file = DEFAULT

make_subdirs = 1

save_lone_pairs = 0

fit_points_file = fit_pts.mol2

read_fitpts = 0

FLAGS

internal_ligand_h_bonds = 0

flip_free_corners = 0

match_ring_templates = 0

flip_amide_bonds = 0
flip_planar_n = 1 flip_ring_NRR flip_ring_NHR
flip_pyramidal_n = 0
rotate_carboxylic_oh = flip
use_tordist = 1
postprocess_bonds = 1
rotatable_bond_override_file = DEFAULT
solvate_all = 1

TERMINATION

early_termination = 0
n_top_solutions = 3
rms_tolerance = 1.5

CONSTRAINTS

force_constraints = 0

COVALENT BONDING

covalent = 0

SAVE OPTIONS

save_score_in_file = 1
save_protein_torsions = 1
concatenated_output =
/home/qthemans/Documents/phgdh/Docking/Ppyr_fpyr/MonoA/Results.mol2
clean_up_option delete_all_solutions
clean_up_option delete_redundant_log_files
clean_up_option delete_empty_directories
output_file_format = MOL2

FITNESS FUNCTION SETTINGS

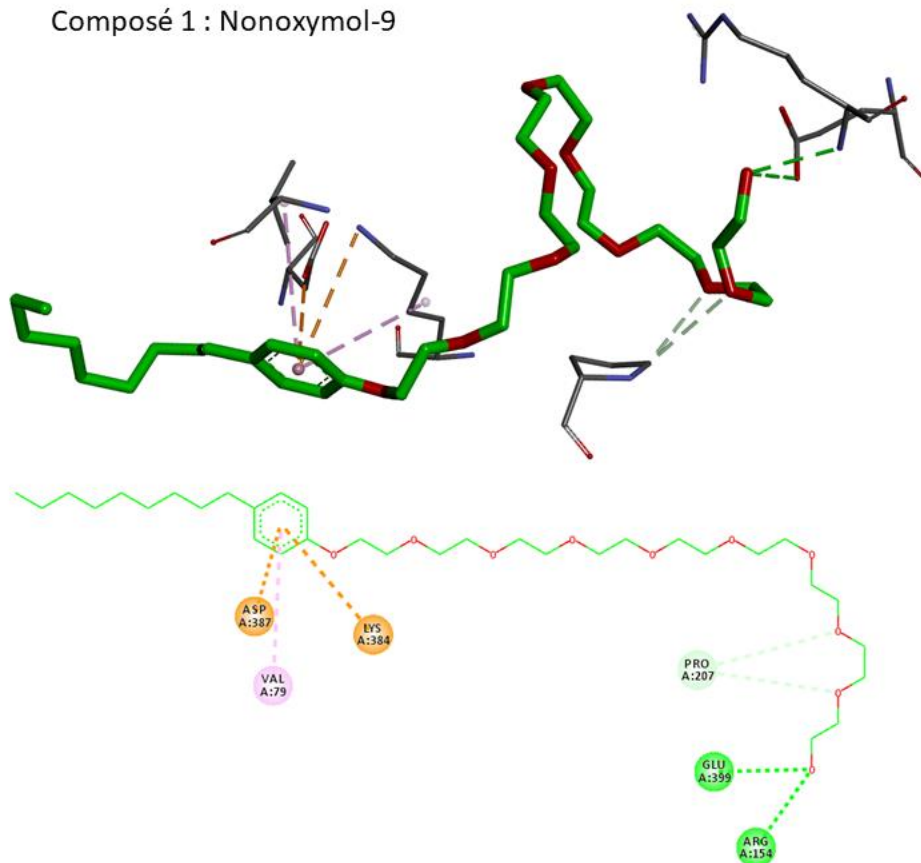
initial_virtual_pt_match_max = 3
relative_ligand_energy = 1
gold_fitfunc_path = plp
score_param_file = DEFAULT

PROTEIN DATA

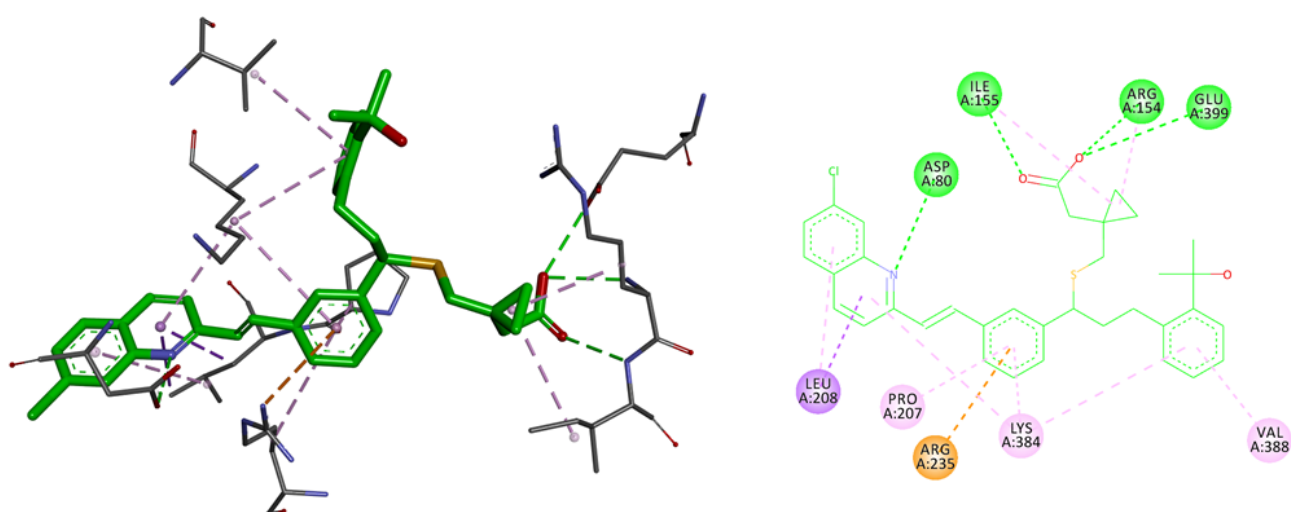
protein_datafile =
/home/qthemans/Documents/phgdh/Docking/Ppyr_fpyr/Tetra_Hybrid_prep_minimize
d_noligands_MonoA.mol2

Annexe C : Mode de liaison des composés identifiés par criblage des sites actifs. Ceux-ci sont présentés par ordre croissant (les composés 6 à 10 ne se liant pas dans la poche ne sont pas présentés).

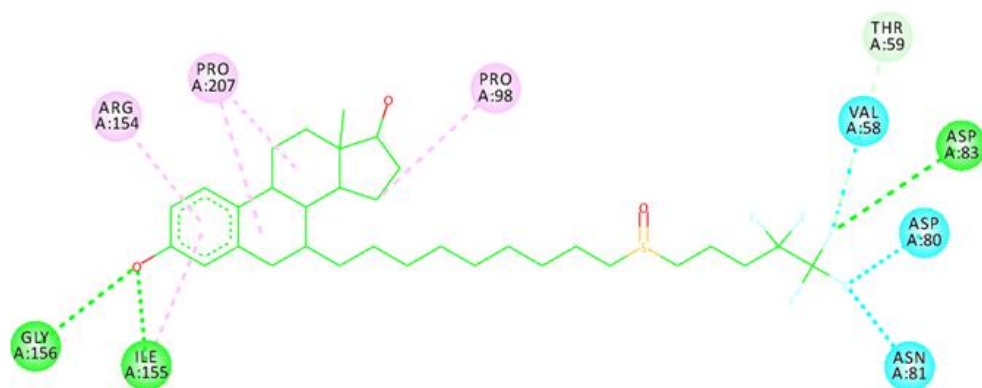
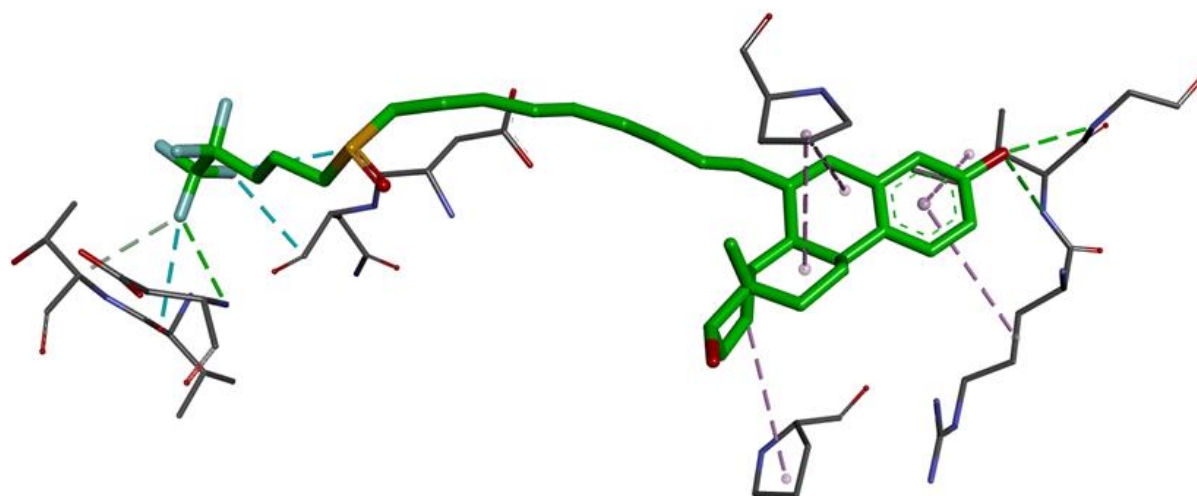
Composé 1 : Nonoxymol-9



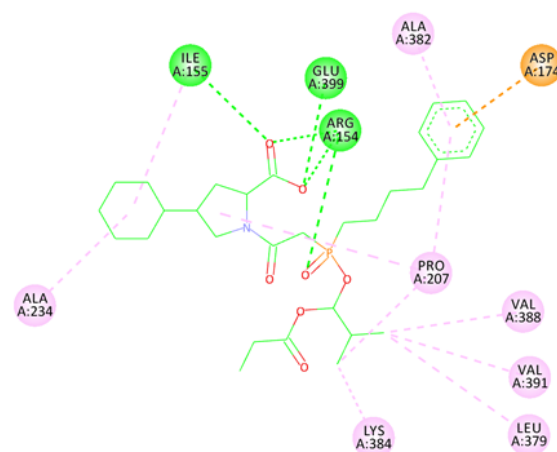
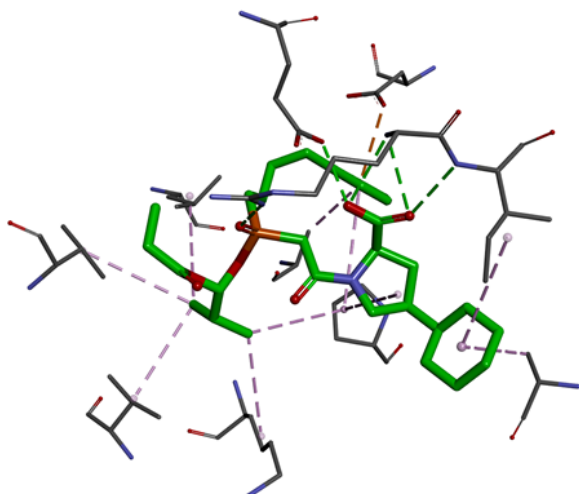
Composé 2 : Montelukast



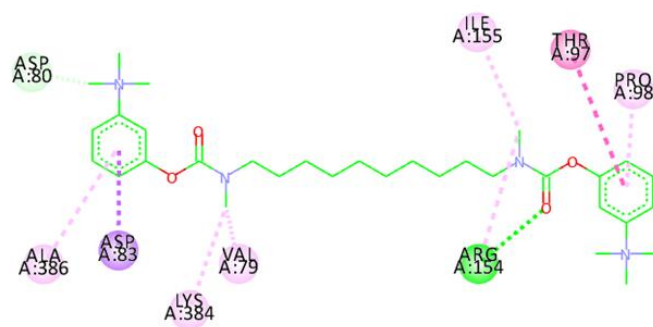
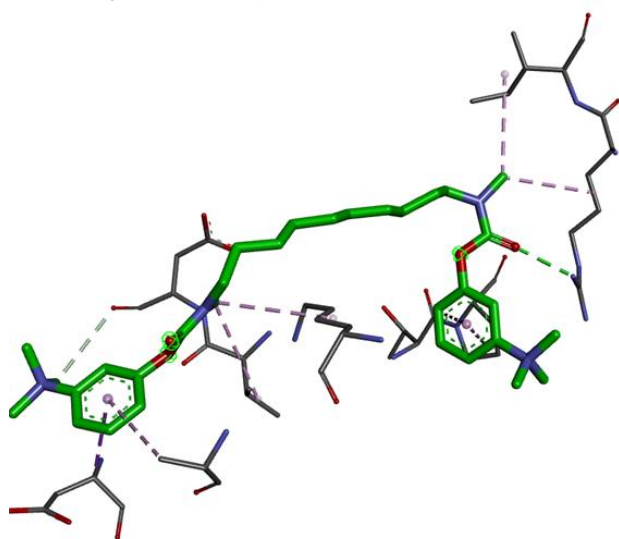
Composé 3 : Fulvestrant



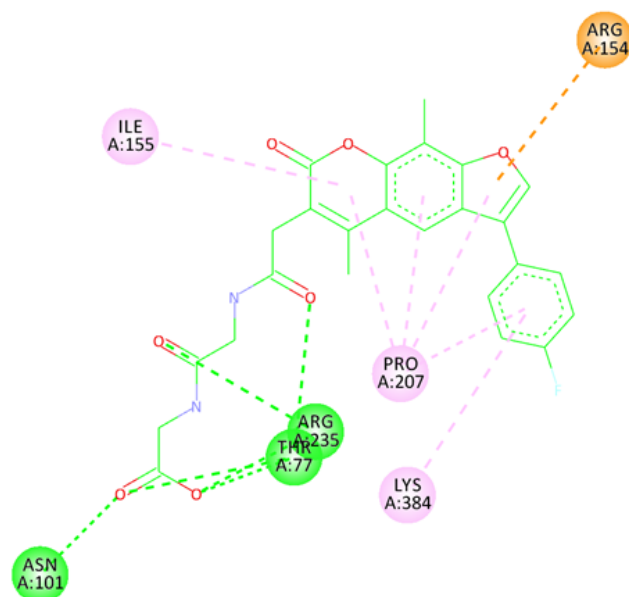
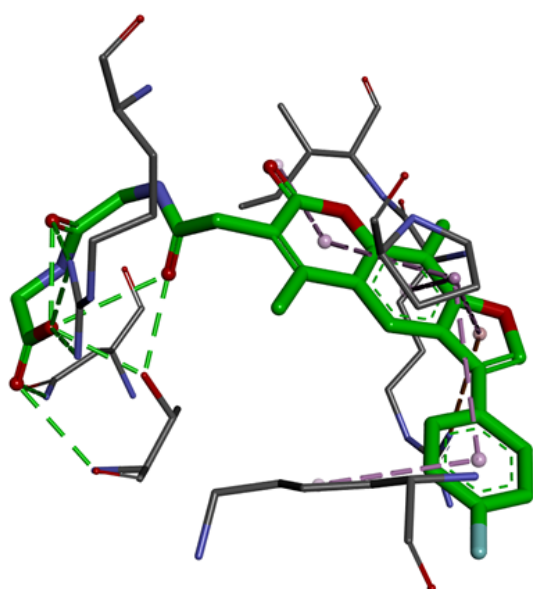
Composé 4 : Fosinopril



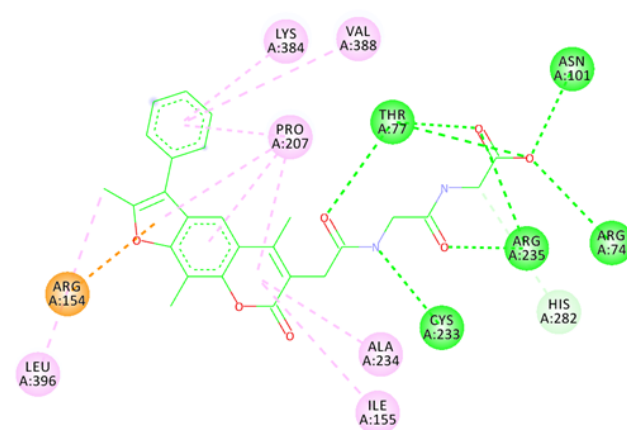
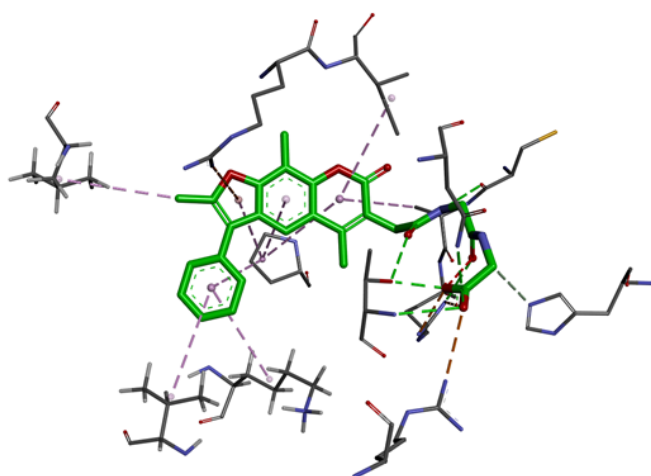
Composé 5 : Demecarium



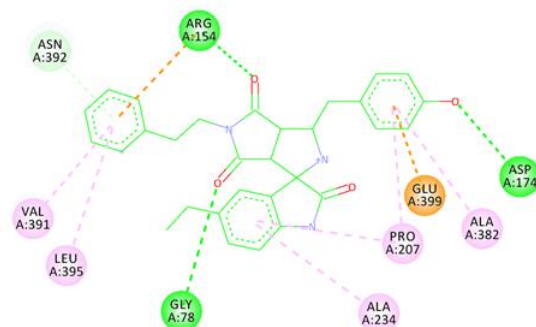
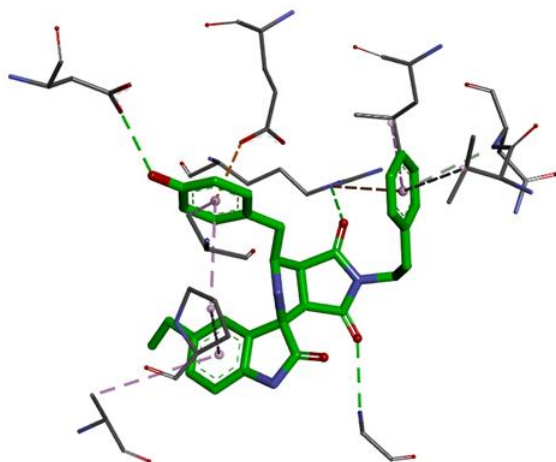
Composé 11 : 2-[[2-[[2-[3-(4-fluorophenyl)-7-keto-5,9-dimethyl-furo[3,2-g]chromen-6-yl)acetyl]amino]acetyl]amino]



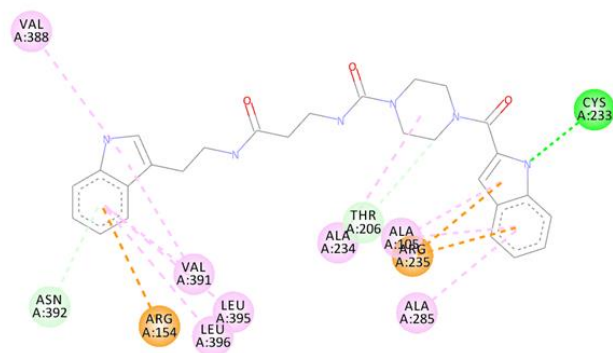
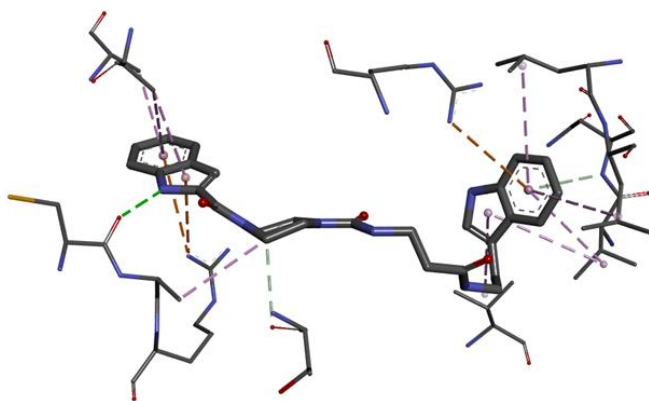
Composé 12 : 2-[[2-[[2-(7-keto-2,5,9-trimethyl-3-phenyl-furo[3,2-g]chromen-6-yl)acetyl]amino]acetyl]amino]acetate



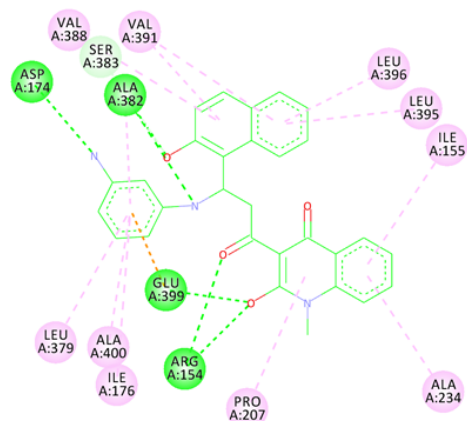
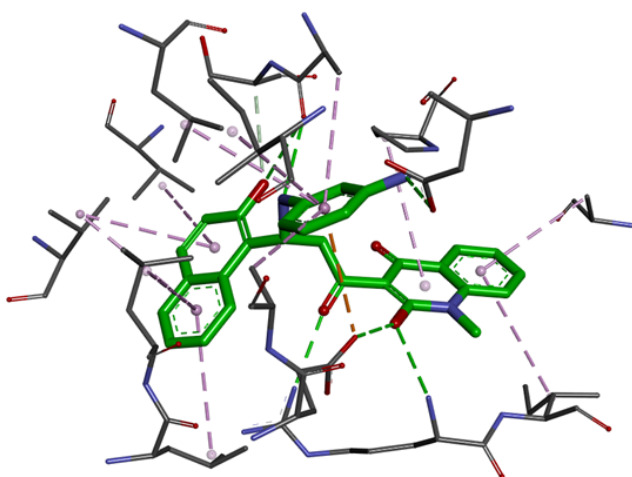
Composé 13 : (1S,3R,3aS,6aS)-5'-ethyl-1-[(4-hydroxyphenyl)methyl]-5-phenethyl-spiro[1,2,3a,6a-tetrahydropyrrolo[4



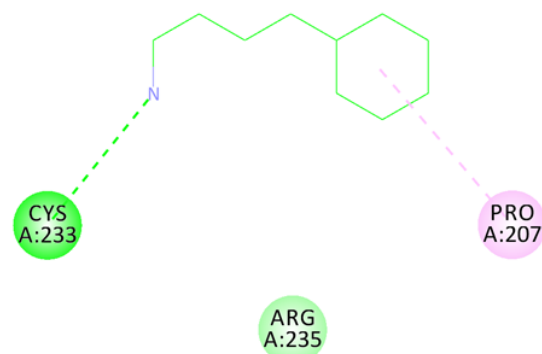
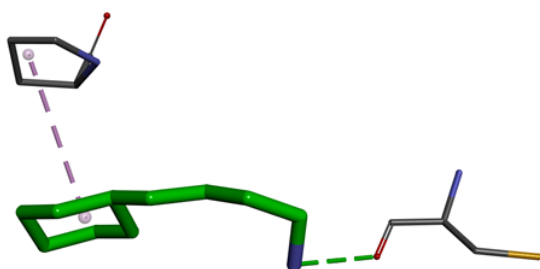
Composé 14 : 4-(1H-indole-2-carbonyl)-N-[3-[2-(1H-indol-3-yl)ethylamino]-3-oxo-propyl]piperazine-1-carboxamide



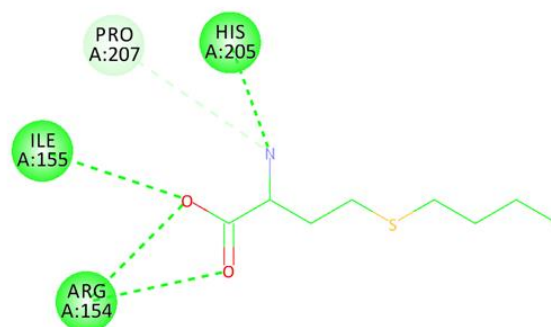
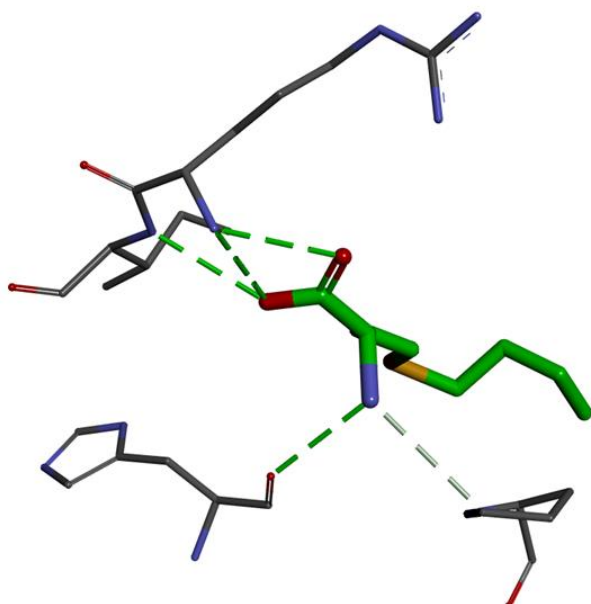
Composé 15 : 3-[3-(3-aminophenyl)amino-3-(2-hydroxy-1-naphthyl)-propanoyl]-4-hydroxy-1-methyl-quinolin-2-one



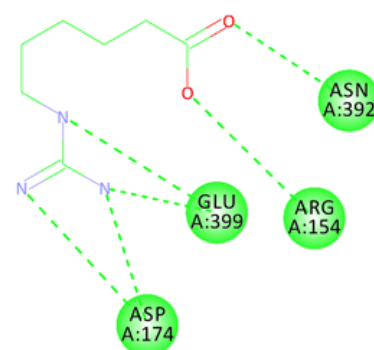
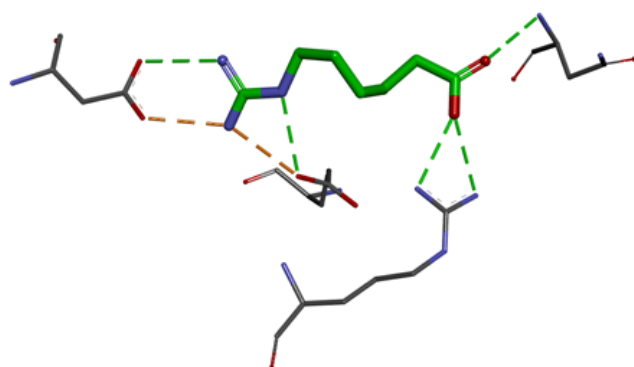
Composé 16 : 4-Cyclohexyl-butylamine



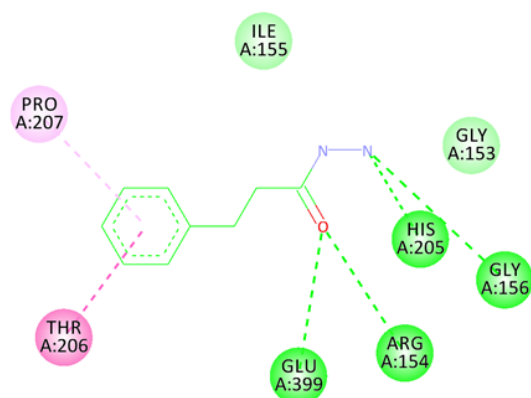
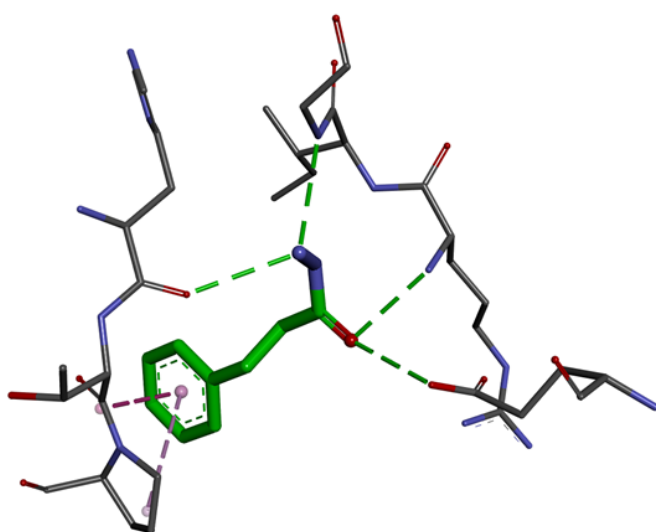
Composé 17 : D,L-BUTHIONINE



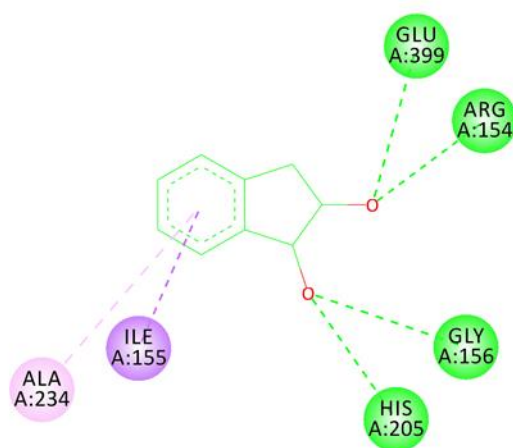
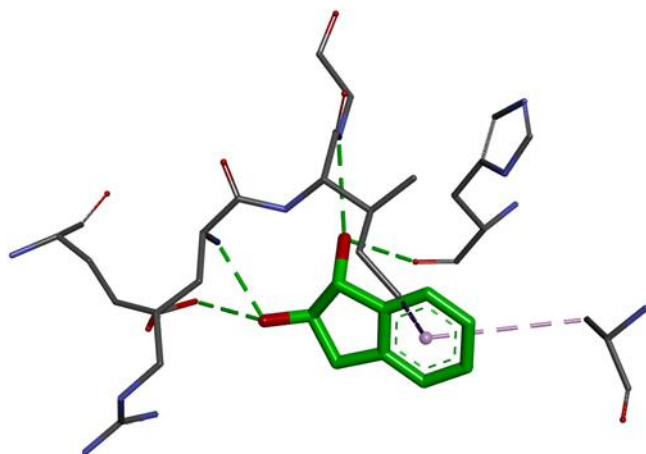
Composé 18 : Acide 6-Guanidinohexanoïque



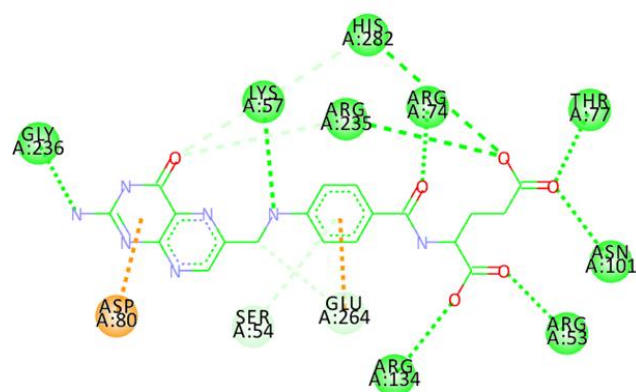
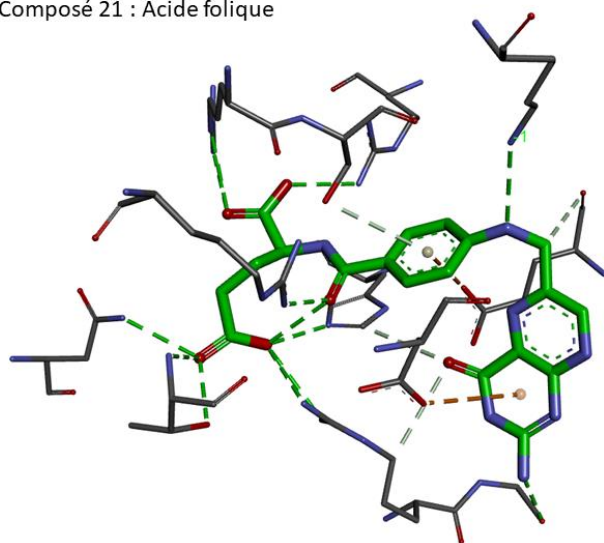
Composé 19 : 3-phenylpropanohydrazide



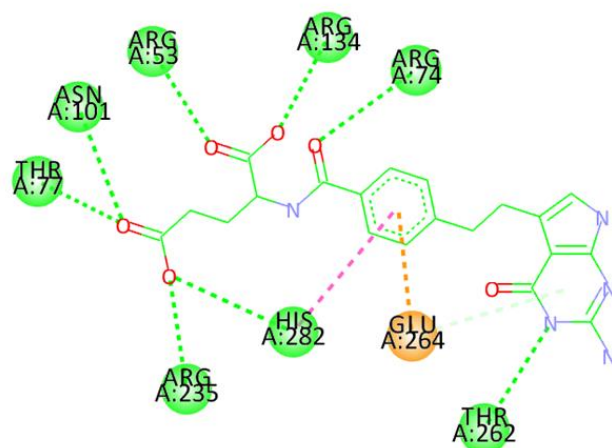
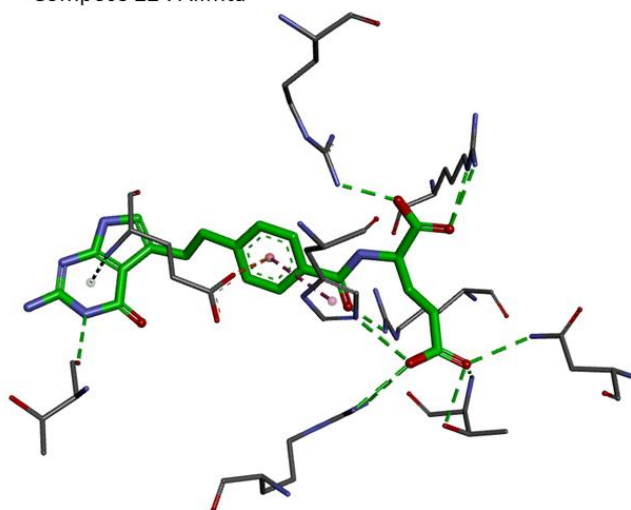
Composé 20 : 1,2-Dihydroxyindan



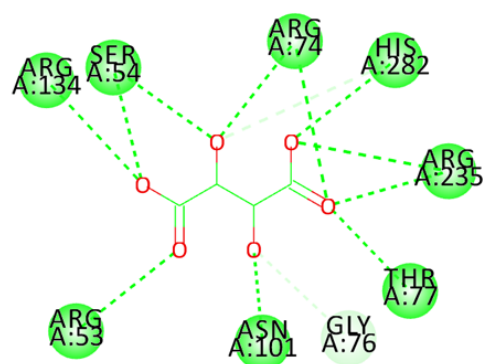
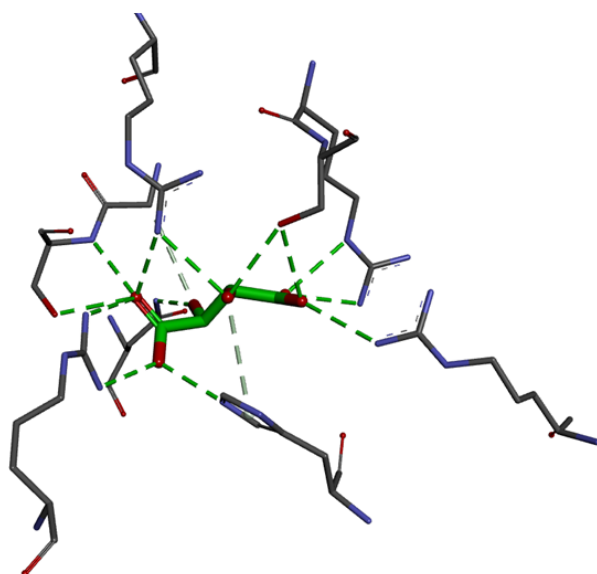
Composé 21 : Acide folique



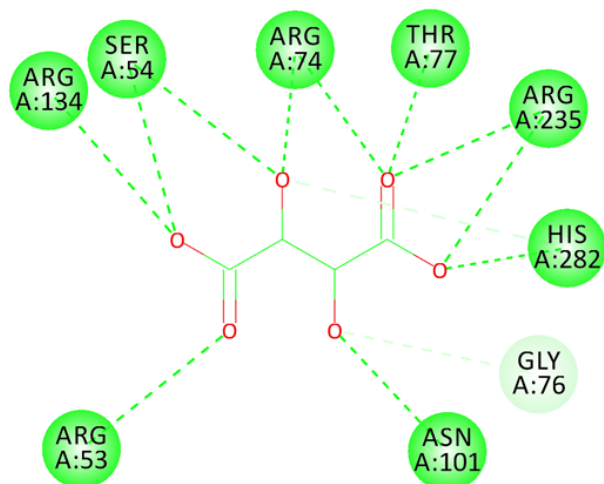
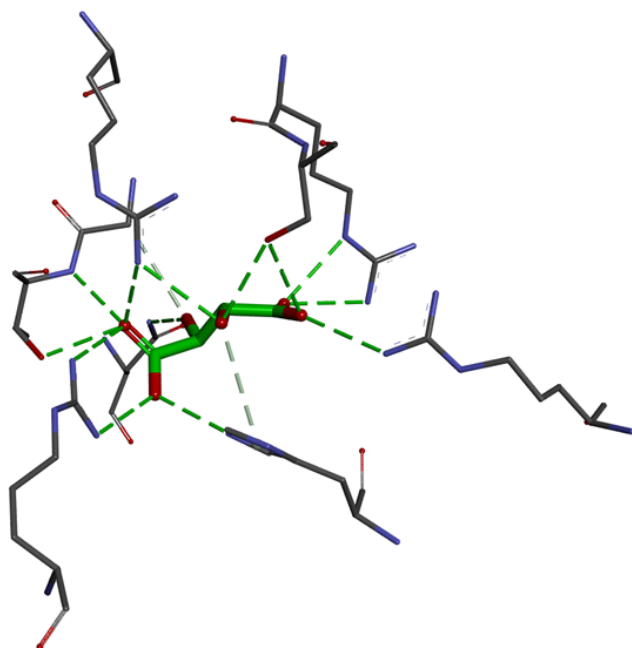
Composé 22 : Alimta



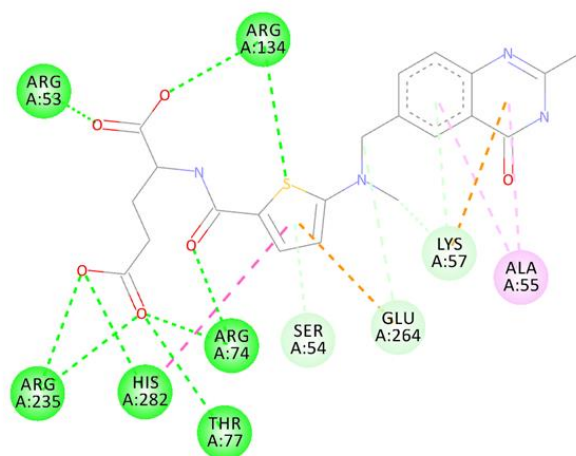
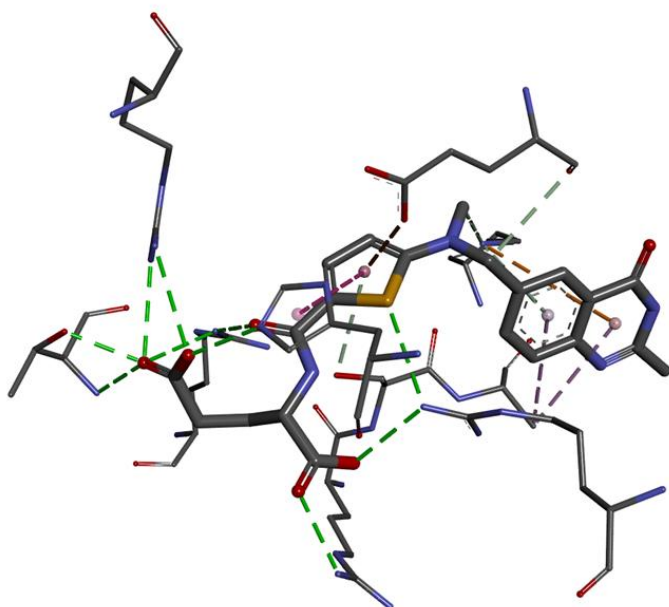
Composé 23 : Acide D-tartrique



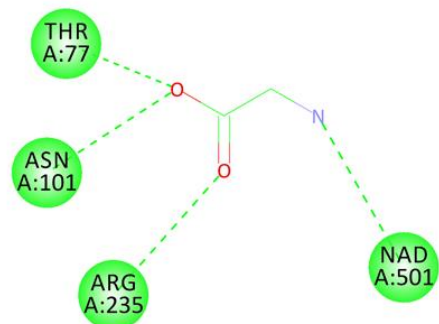
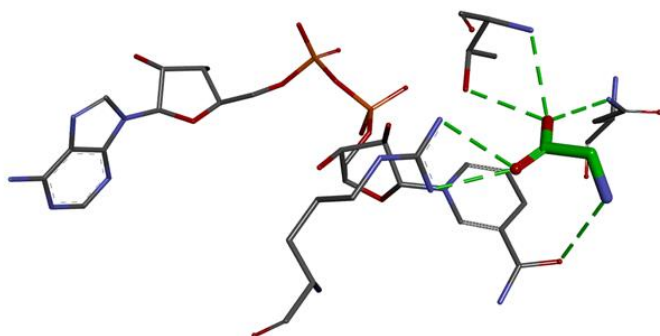
Composé 24 : Acide L-tartrique



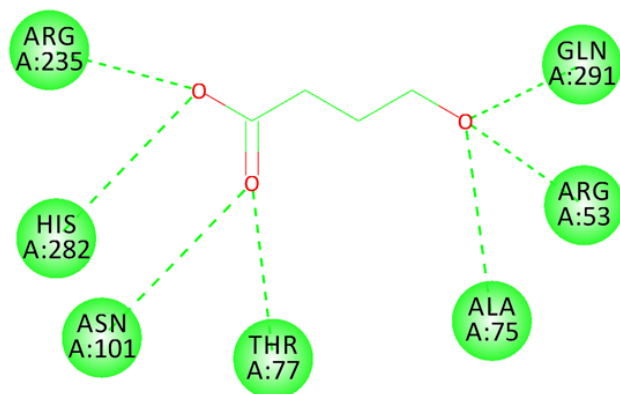
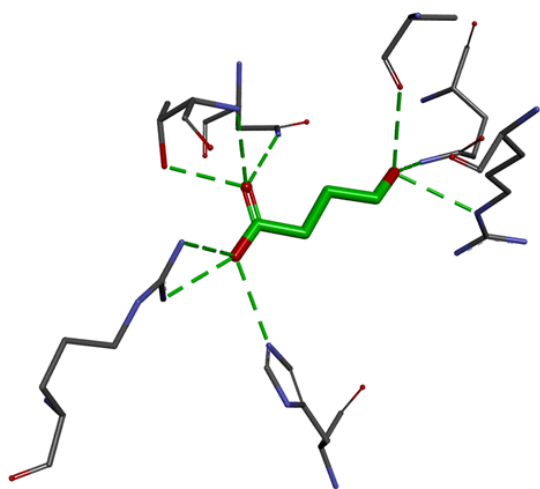
Composé 25 : Raltitrexed



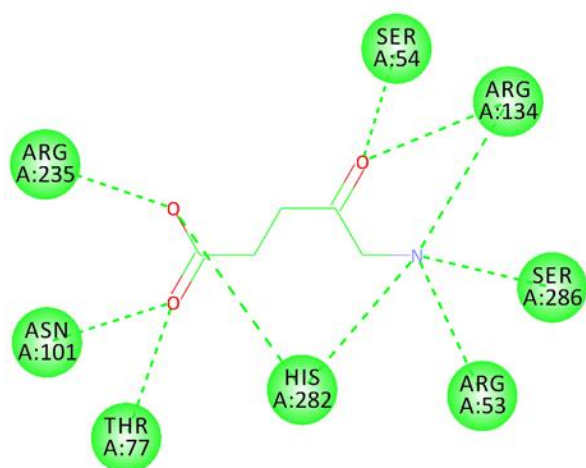
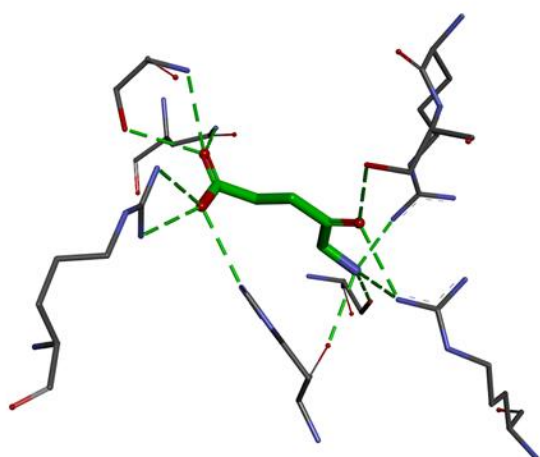
Composé 26 : Glycine



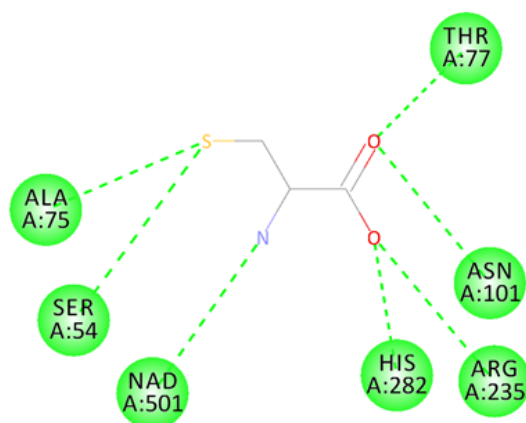
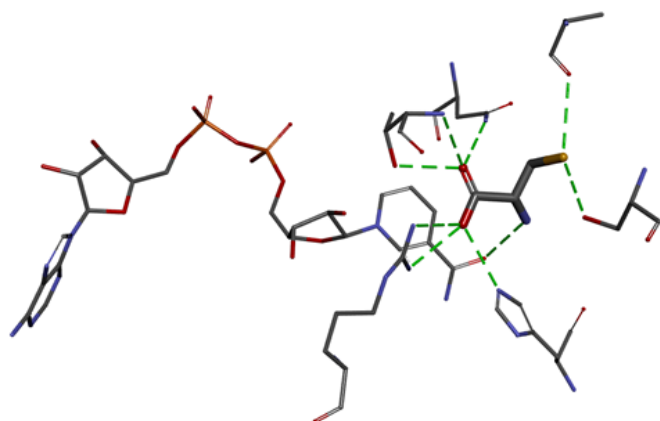
Composé 27 : Oxybate de sodium



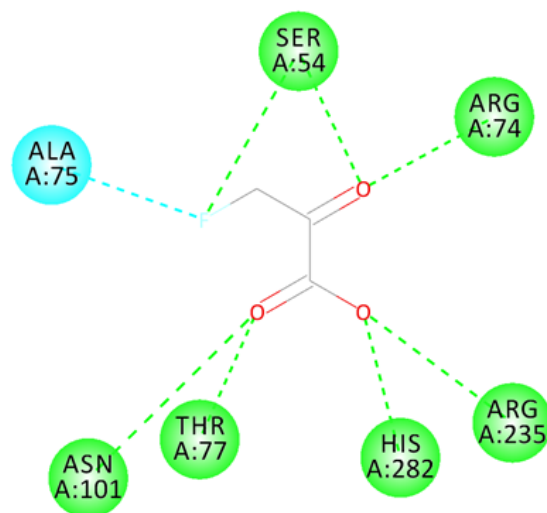
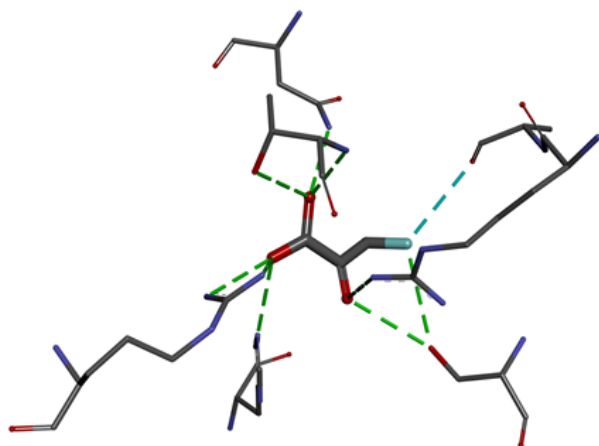
Composé 28 : Acide 5-Aminolevulinique



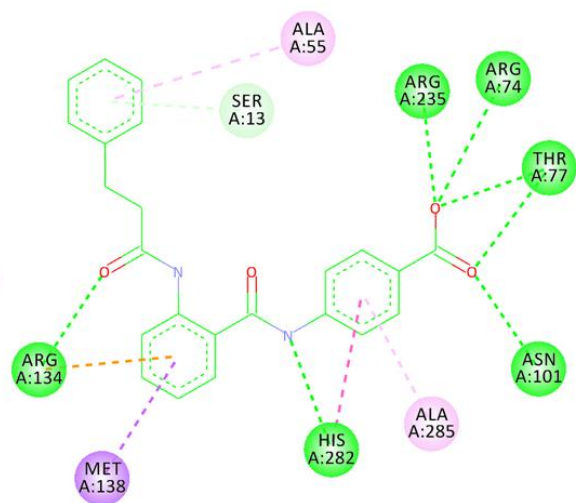
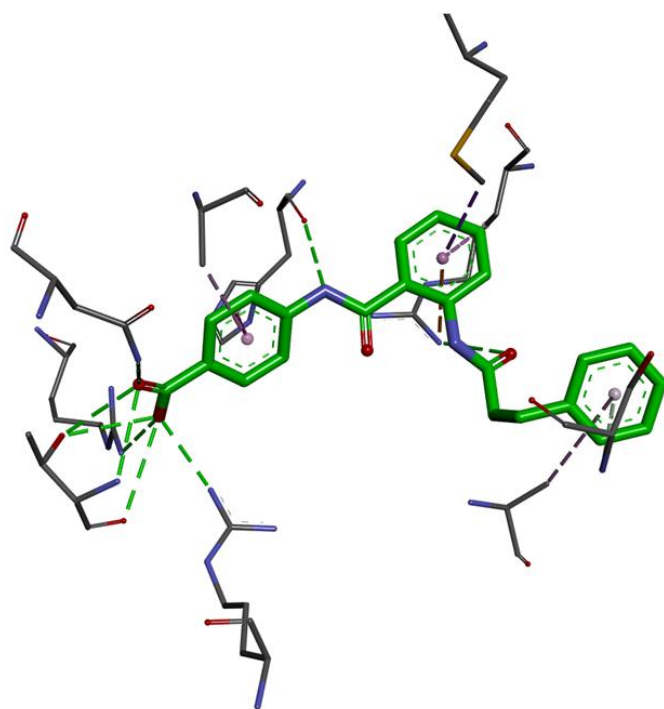
Composé 29 : L-cystéine



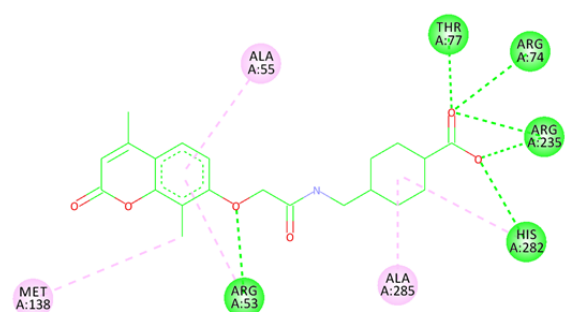
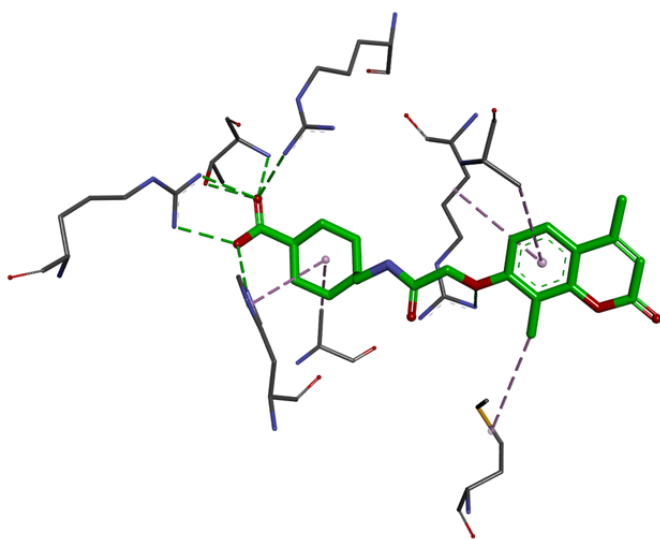
Composé 30 : Acide fluoropyruvique



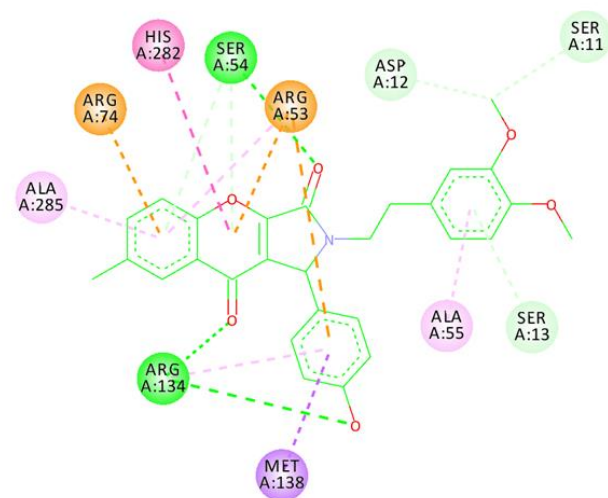
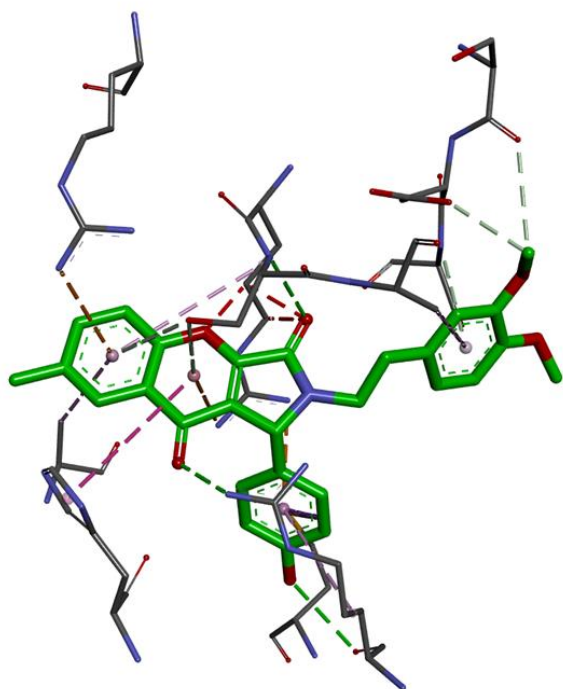
Composé 31 : 4-[2-(3-phenylpropanoylamino)benzoyl]am



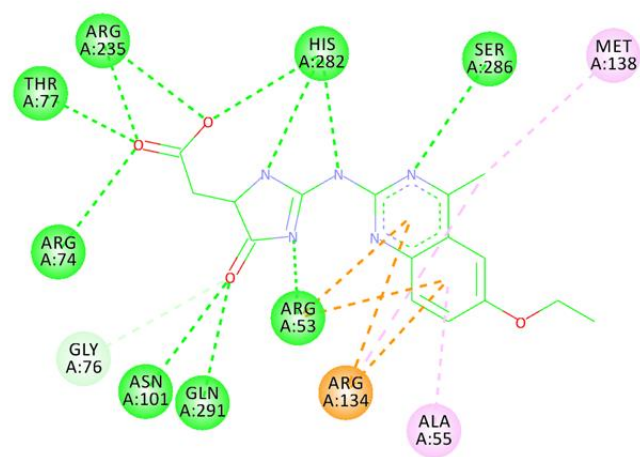
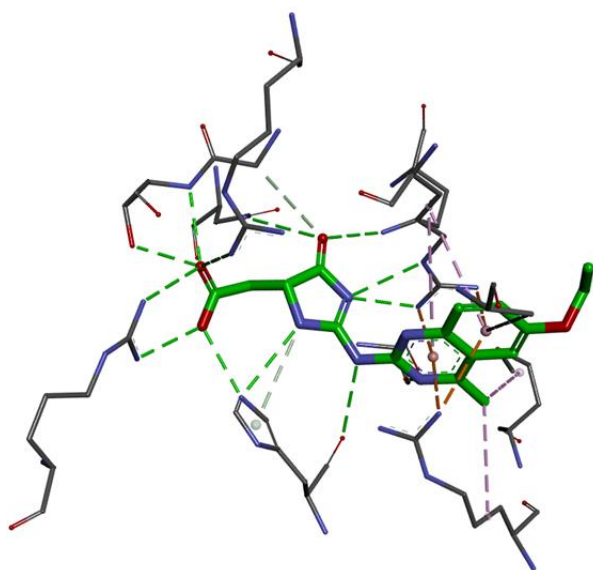
Composé 32 : 4-[[[2-(2-keto-4,8-dimethyl-chromen-7-yl)oxyacetyl]amino]methyl]cyclohexanecarboxylate



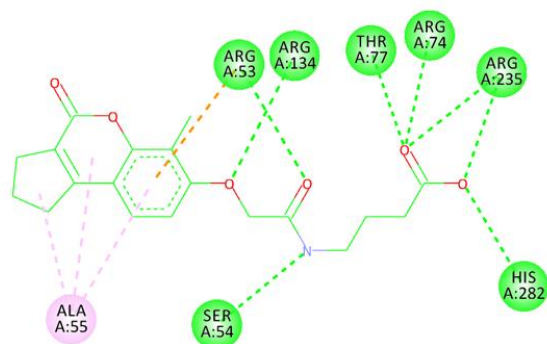
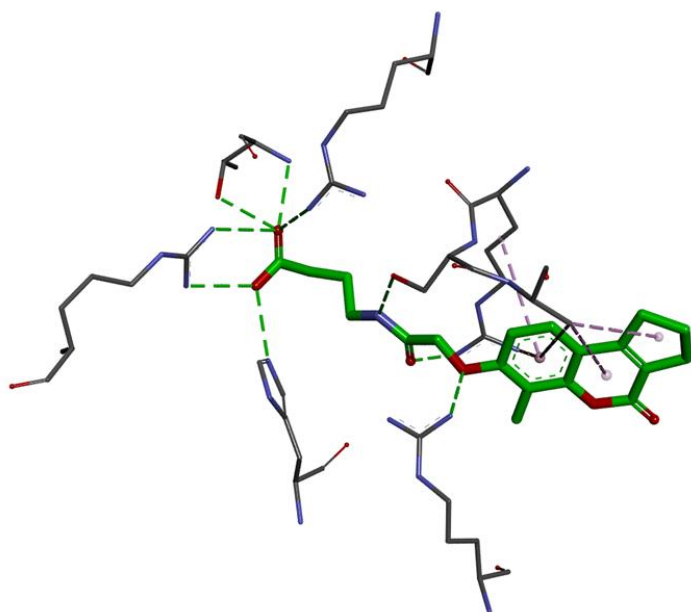
Composé 33 : (1S)-2-homoveratryl-1-(4-hydroxyphenyl)-7-methyl-1H-chromeno[2,3-c]pyrrole-3,9-quinone



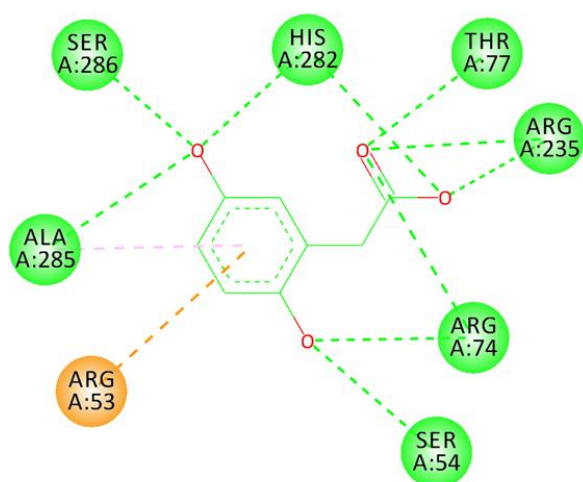
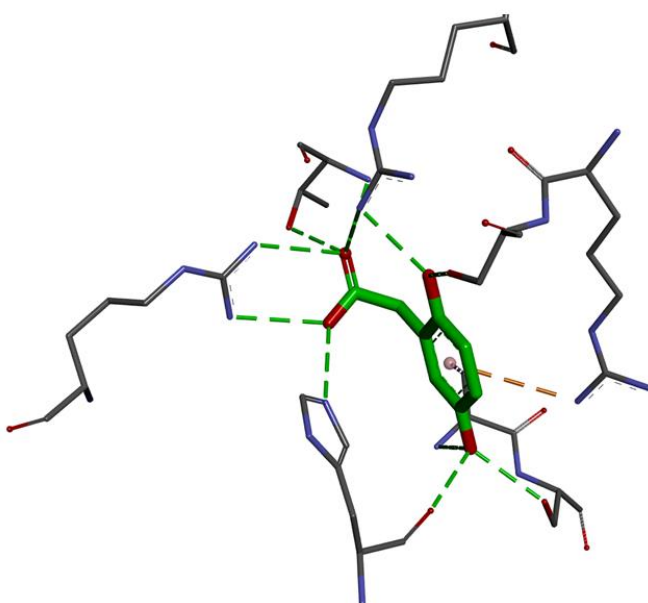
Composé 34 : 2-[(4R)-2-[(6-ethoxy-4-methyl-quinazolin-2-yl)amino]-5-oxo-1,4-dihydroimidazol-4-yl]acetic



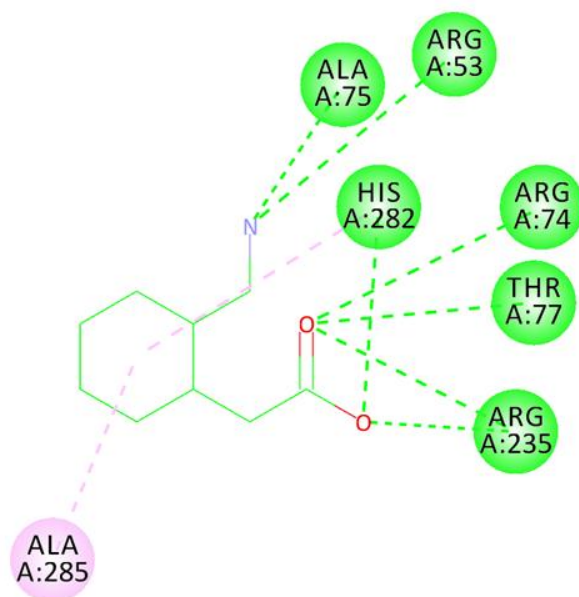
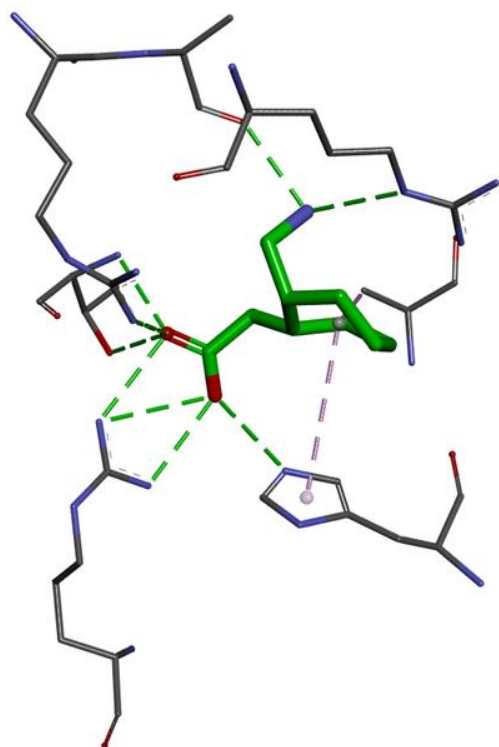
Composé 35 : 4-[[2-[[4-keto-6-methyl-2,3-dihydro-1H-cyclopenta[c]chromen-7-yl]oxy]acetyl]amino]butyrate



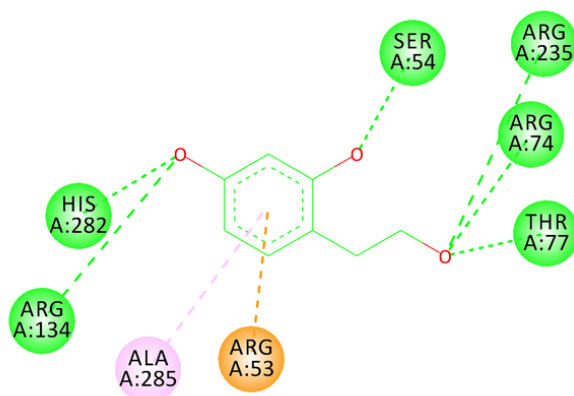
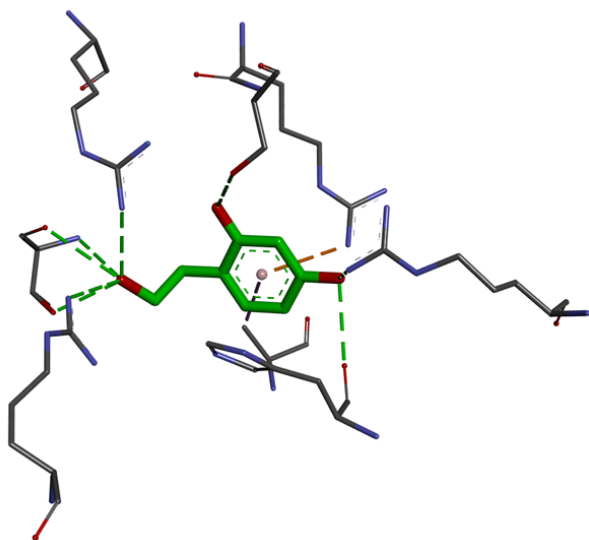
Composé 36 : Acide homogentisique



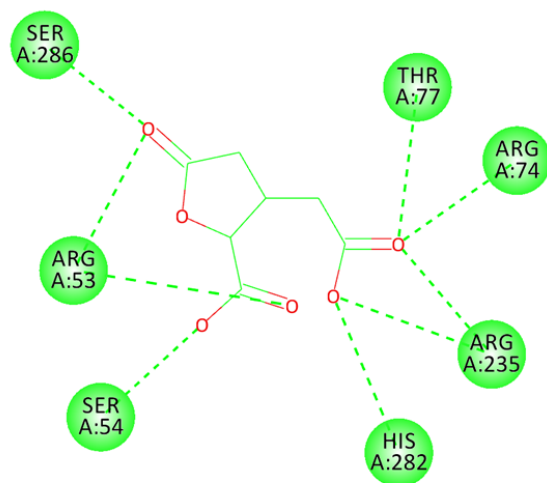
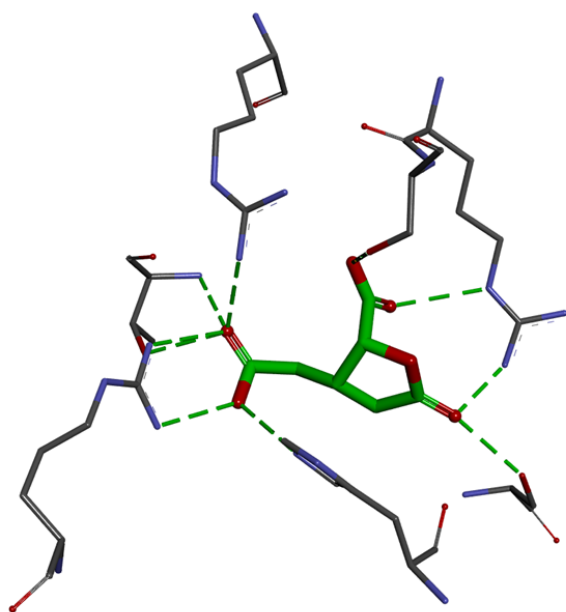
Composé 37 : Gabapentin



Composé 38 : 4-(2-hydroxyethyl)benzene-1,3-diol



Composé 39 : 3-(carboxymethyl)-5-oxo-tetrahydrofuran-2-carboxylique



Composé 40 : Danshensu de sodium

